POWER SYSTEM OPTIMISATION:

Deterministic and Stochastic Annual Scheduling

A thesis

presented for the degree of

Doctor of Philosophy in Electrical Engineering

in the

University of Canterbury,

Christchurch, New Zealand

by

T. S. Halliburton B.E. (Hons)

1983

CONTENTS

PAGE

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF SYMBOLS

$\underset{\sim}{d}_i$ - search direction vector.

D - electrical energy demand (MWh).

DY - directional derivative.

ECC - loss coefficient, matrix, for D.C. loadflow method.

EC0 - loss coefficient, vector.

E00 - loss coefficient, scalar.

erf(.) - error function.

g_v - gradient.

h_j - duration of load class j (hours).

$\underset{\sim}{G}$ - vector of generator energy outputs.

G_v - generation from a river system (MWh).

$\underline{G}$ - minimum generation (MWh).

$\bar{G}$ - maximum generation (MWh).

H^i - Hamiltonian.

I_v - inflow (MWh).

i - inflow, one reservoir model, (GWh).

$\bar{i}$ - mean inflow (GWh).

j - load class index.

k - conjugate direction index.
- gain, linear feedback model.

$k_{q/q+1}$ - conversion factor for water, from equivalent energy value in reservoir q to value in reservoir q+1.

K_v - water flow to energy conversion factor (MWh/cumec).

ℓ - number of load classes.
- index for inflow probability distribution points.

L(t,j) - transmission loss.

$L(x,\lambda)$ - Lagrangean, a function of x and Lagrange multiplier λ.

m - number of reservoirs.

MX_v - range of releases possible from reservoir (MWh).

n - number of reservoirs plus number of thermal stations.

p - number of run-of-river stations
- probability.

R_v - reservoir water release (MWh).

RR - run-of-river stations index.

$\underline{R}_v$ - minimum release (MWh).

$\bar{R}_v$ - maximum release (MWh).

S_v - spill (MWh).

RSC - restart parameter.

SD(i) - element of search direction vector.

SZ - linear search step size.

t - time interval index.

t_f - final time.

T_v - tributary inflow (MWh).

u - release, one reservoir model, (GWh).

$\bar{u}$ - release at mean storage, one reservoir model, (GWh).

U_v - unconstrained release.

v - reservoir or thermal station index.

W_v - penalty function weight, on state variables.

$W_{(f)v}$ - penalty function weight on final time target state.

WU_v - penalty function weight on control variable.

x - reservoir storage, one reservoir models, (GWh).

$\bar{x}$ - mean storage, one reservoir models (GWh).

X_v - reservoir storage (MWh).

$\underline{X}_v$ - minimum storage (MWh).

$\bar{X}_v$ - maximum storage (MWh).

XLS - linear search accuracy parameter.

α - probability of maximum storage.

δ - delta function (impulse).

λ - Lagrange multiplier.

λ_v - incremental cost of thermal fuel.

μ - mean of Gaussian.

ρ_v - costate.

σ - unit step function
- standard deviation.

σ_x - standard deviation of storage.

σ_i - standard deviation of inflows.

ABSTRACT

The optimisation of hydro-thermal power systems over a one year horizon is investigated. The objective is to minimise fuel costs by appropriate scheduling of releases from large hydro storage lakes. Fuel cost is the principle avoidable expense in operating such a system as hydro station costs are independent of power generated.

A deterministic model is described. It represents all stations fed by each reservoir as one equivalent station. All generation and storage constraints are handled. Transmission losses are determined by a simple d.c. load flow loss coefficients method.

A Hamiltonian approach is used to convert this model to a static problem which we can solve by a non-linear programming method. A conjugate gradients algorithm was developed, with emphasis on robustness and accuracy, which successfully solved a number of variants of the problem. As an example, the New Zealand system is solved. It includes eight reservoirs, six thermal stations, and the transmission constraint on a D.C. link.

A one reservoir equivalent of this system is used to study various stochastic inflow effects. Stochastic dynamic programming is shown to give quite different results to deterministic methods. A linear feedback algorithm representing storage probability distributions by a Gaussian plus an impulse is tested. This sub-optimal method does not suffer from the dimensionality barrier of D.P. methods. It is thought to have the potential to provide a practical stochastic scheduling method, for the multi-reservoir case.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Dr. H. R. Sirisena, for his guidance over the five years of research for this thesis. His great patience, and ability to generate enthusiasm when my research has seemed fruitless have been consistent throughout this time.

Dr. J. Lermit of New Zealand Electricity made a very valuable contribution in the early stages of my work with data for lake inflows etc. Mr. M. Turner, also of N.Z.E., has assisted considerably with further system data.

All the computer results in this thesis were obtained on the University of Canterbury, Electrical Engineering Department's VAX 11/780 computer. Without this machine, this thesis would not have been possible in its present form. I am extremely grateful to Bill Kennedy for his tremendous efforts with this computer, both in its initial installation and in ensuring that it ran smoothly.

I am indebted to the University Grants Committee and New Zealand Electricity for financial support during my studies.

Finally, I wish to thank Gordon Tillier of the N.Z.E., Christchurch, draughting office and my typist, Beryl Nottingham, for their efforts in the preparation of this thesis.

CHAPTER 1

INTRODUCTION

This thesis presents some efforts to solve the problem of optimal annual scheduling of hydro-thermal power system operations. The objective of such an optimisation is to minimise the cost of fuel burnt in thermal power stations, by determining appropriate weekly water releases from large long-term hydro electric storage reservoirs. The cost per MWh generated at thermal stations varies greatly from those stations designed for supply of short peak loads only, down to those intended for base load generation. Hence the algorithm must try to ensure that the more costly thermal stations are used as little as possible, that reservoir water is not wasted unnecessarily by spilling, and take into account transmission line losses, and the various physical system constraints.

A variety of optimisation problems, with differing time scales, can be formulated for power system operations. With a time scale of many years, is the optimal system expansion problem. A one week time horizon problem might be solved to find detailed generation schedules, within the constraints of desired total weekly generations determined by the annual problem. An economic dispatch program could be re-run every 20 minutes or so to determine optimal machine loadings, minimise line losses, and check on system security (e.g. line limits, equipment ratings). This is a static problem, whereas all the others are dynamic.

The annual scheduling problem was chosen for this thesis as it is an especially interesting optimisation problem, not due to a desire to solve this particular problem for its own sake. It involves reservoir inflows which are variable, with considerable uncertainty over the one year time scale. This gives opportunities to investigate stochastic control systems. The problem is of high dimension - many reservoirs to schedule over each time interval, and with many time intervals. (A 312 dimensional problem is solved in Chapters 5 and 6). The power system data required is readily available. By optimising a large, real system it was hoped that the development of yet another optimisation algorithm of limited usefulness would be avoided. A good example of what we wished to avoid is briefly described by Anonymous (1972).

The structure of the problem is also interesting in that it consists of a number of subsystems with only limited coupling. Each river system with a chain of hydro stations along it, or in some cases only one station, forms a subsystem. Coupling is largely restricted to the requirement that total generation equals demand.

Research efforts began on a deterministic model, with a reliable solution method. A model of the New Zealand system was solved with 8 large reservoirs (two of which are in series) and 6 thermal stations. Then attention moved to stochastic models. Techniques are less well developed in this area. So a simple model was used with attention concentrating on the aspects of the hydro-thermal problem which can not be handled satisfactorily by established procedures.

The thesis proceeds as follows:-

Chapter 2 first describes previous work on the annual scheduling of the New Zealand power system. None of these methods appear to have been used to schedule reservoir releases, in practice. Some representative examples of deterministic and stochastic models are discussed to help identify the state of the art, and to give a basis for the evaluation of our work. A decomposition scheme which we experimented with, unsuccessfully, is described.

Chapter 3 gives the deterministic modelling technique in general terms. The problem is formulated as a large non-linear program to be solved by the unconstrained, static hill-climbing algorithm of Chapter 4. The various constraint enforcement strategies and conversion from a dynamic to a static problem are described.

Chapter 4 describes the development of a generalised conjugate gradients algorithm, "CGRADS". A Fletcher and Reeves version was used first. Problems encountered and the steps taken leading to the final form of "CGRADS" are described. This was necessary to obtain reliable solutions to our very difficult problem.

Chapter 5 applies the modelling technique of Chapter 3 to the New Zealand system, and gives detailed model data. We consider this model is as accurate as possible, with the data available.

In Chapter 6, solutions to the New Zealand model are presented in detail for one case. The validity of the results is discussed. The various constraint enforcement methods experimented with are compared.

In Chapter 7 we first go through various aspects of reservoir inflow modelling, to help understand the nature of the stochastic beast. Then stochastic and deterministic dynamic program results are given for a one reservoir model, followed by a linear feedback method. This method is sub-optimal, but feasible for a multi-reservoir problem, unlike the dynamic programming methods.

Chapter 8 develops the linear feedback method further by using a non-gaussian distribution for water storages. This method, if developed further could offer a practicable solution to the multi-reservoir stochastic problem.

Chapter 9 concludes this thesis with comments on implementation of a scheduling method. Some ideas are given on how the work described here could be extended to produce a computationally feasible, stochastic method which could possibly be a useful, practicable solution to this problem.

All computer results given in this thesis were obtained on a VAX 11/780 computer.

The following papers were published in the course of research for this thesis:

Sirisena, H. R. and Halliburton, T. S., "Long-term optimisation of hydro-thermal power systems by generalised conjugate gradient methods", Optimal Control Applications and Methods, 2, pp351-364, 1981.

Halliburton, T. S. and Sirisena, H. R., "Long-term optimal operation of a power system", I.E.E. Proc., 129ptC, pp185-191, 1982.

Halliburton, T. S. and Sirisena, H. R., "Development of a stochastic optimisation for multi-reservoir scheduling", IEEE Trans Automatic Control, to appear Nov. 1983.

CHAPTER 2

SURVEY OF SOME POWER SYSTEM SCHEDULING METHODS

2.1 INTRODUCTION

This chapter surveys a few examples of reservoir scheduling algorithms to help put the ideas developed in this thesis into perspective. A detailed study of any other method is not made as our method was not inspired by, or developed from any other.

Extensive bibliographies on the topic exist. Rosenthal (1980) gives a (by no means exhaustive) list of 100 papers and classifies them according to their ability to handle:

(i) Multiple reservoirs

(ii) Multiple time periods

(iii) Stochastic inflows

(iv) Non-separable cost functions (F) as a measure of system benefits i.e. $\frac{\partial^2 F}{\partial x_i \, \partial x_j} \neq 0$ for $i \neq j$.

None of the 100 methods listed has all these desirable properties.

Sachdeva (1982) gives a bibliography for reservoir scheduling methods with a variety of time scales. 110 of these refer to the long term problems despite covering only the period 1960 to 1978. El-Hawary and Christensen (1979) describe their own work on shorter time scale problems in detail and are a source of references to literature on power system optimisation generally.

Models developed for the New Zealand system are covered in more detail than others, and are described separately in the next section. Following this, some deterministic and then stochastic models are described. Finally a decomposition method is given in some detail. Considerable effort was expended on its investigation, as it appeared to have some merit. The difficulties leading to this approach being abandoned will be discussed.

2.2 NEW ZEALAND SYSTEM MODELS

No long term scheduling algorithm of the type described in this thesis is in use by New Zealand Electricity, but a number of aspects of the system have been modelled for optimisation purposes.

One of the first (McCool *et al* (1966)) was a simulation (not an optimisation) to study the effect of diverting Tongariro catchment water into Lake Taupo. Three reservoirs were modelled, all South Island storage being lumped into one. Simulations were over one year's operation with a one day time step. Fears had been expressed about the possibility of flooding due to the extra lake inflows. As a result of the simulation, lower lake levels were recommended, thereby reducing the possibility of lake side flooding, despite higher mean flows.

Green (1971a, b) used two dimensional dynamic programming to obtain a given electrical output from the three lower Waitaki stations as economically as possible. This involves minimising spill, operating with the highest possible heads, and running generators at their peak efficiencies as much as possible. A one day horizon and one hour time step were used.

McKerchar (1971, 1975) used deterministic dynamic programming and synthetic streamflows to solve a system consisting of Lakes Pukaki and Tekapo, and assumed quadratic thermal generation costs. He simulated 40 years of streamflows, then used linear regression to obtain reservoir release as a function of the storage at the beginning of a time interval. A one month time step was used.

Lusk (1972) describes the "Basic Rule Curve" method which has been used by New Zealand Electricity (NZE). This involves producing a diagram of storages against time. It indicates the storage level at which maximum thermal generation has to be brought into use if an energy deficiency is to be averted under the worst possible inflow conditions. This is a security, not an optimisation device, however. It is used to draw up guidelines which do take some account of economic operation. Lusk(1976) proposes a trajectory method, based on that of Electricite de France (EDF) (Daellenbach and Read (1976)). Trial and error is used to locate reservoir storage trajectories which give incremental costs of thermal energy as near to equal as possible for each month of the year.

Boshier and Lermit (1977) produced a network flow algorithm intended primarily for estimating marginal energy costs. It involved weekly time steps, seven reservoirs and the D.C. link constraints. Each time interval was split into five load classes derived from the load duration curve, as for the method described in Chapter 3. A 1½ year time horizon was used.

Daellenbach (1979) proposed a stochastic dynamic programming method decomposing the system into separate North and South Island subsystems, joined by the D.C. link. A pricing mechanism would co-ordinate the subsystems.

Reed (1979) adapts the E.D.F. trajectory method to give a decomposition approach. He describes a detailed model involving losses and transmission constraints. It is suitable for short term scheduling of a chain of stations, such as on the Waikato, and for the long term problem. An example problem is solved. The model used includes six reservoirs and has a one week time step. The D.C. link is the only transmission constraint handled. The possibility of a stochastic solution using the work of Rockafellar and Wets (1976) on optimal recourse is considered. The same decomposition method is used as in the deterministic case, but with a different hydro subproblem.

N.Z.E. have developed a one reservoir stochastic method, but do not appear to have published a description. It is based on the work of Stage and Larson (1961), which is effectively a stochastic dynamic programming method. It has been used to help estimate thermal fuel requirements.

2.3 DETERMINISTIC MODELS

Models which do not include the stochastic nature of inflows are considered here. The first group of five models were not tested on realistic systems - only on simple contrived systems. The second group of four were developed to produce schedules for real systems or at least for realistically large problems.

Agarwal and Nagrath (1972) solved a system of two hydro stations and two quadratic cost thermal stations, over 12 one month periods. Three different optimisation methods were compared.

Fults *et al* (1976) examined a four reservoir multi-user system involving hydro electricity generation, flood control, irrigation, city water supply and navigation uses. One month time intervals, on a one year horizon were used with incremental dynamic programming. The successive approximations method involved optimising one reservoir at a time, keeping the strategy for the other three fixed. The choice of the initial strategy was found to be crucial, and some convergence problems were encountered.

Saha and Kharparde (1978) looked at two hydro stations, two quadratic cost thermal stations, and used 12 one month intervals. Feasible directions and conjugate directions methods were compared.

Kumar *et al* (1979) used decomposition in time, applied the method of multipliers to handle subproblem constraints, and were able to solve a two reservoir problem in six minutes on an IBM360 computer.

Soares *et al* (1980) handled a stochastic load function, incorporating a penalty in the cost function for shortfalls or excesses. Penalty costs represented the cost of buying in energy or the value of sales of surpluses. By forming an additively separable Lagrangian, decomposition in time and space is possible with this method. A system of four hydro stations, two quadratic cost thermal stations and 12 time intervals is solved.

The techniques described in these papers do not seem of great value in attempting to solve a real problem. Convergence of the algorithms is likely to be much more difficult in a larger problem involving a wide range of reservoir sizes, and many more time intervals. Quadratic thermal station costs are also a common feature of these methods. This gives a conveniently smooth, differentiable cost function, but is unrealistic for the long term problem. Many machines at a number of thermal stations will usually be involved in practice, and a piecewise linear cost function may be more appropriate. Our experience indicates that optimisation algorithms that are effective on small problems with smooth contours are not necessarily useful on large problems with awkward but realistic features. Some work with more realistic system models follows.

The U.S. Pacific Northwest system optimisation is described by Hicks *et al* (1974) and Gagnon *et al* (1974). This is a purely hydro system, so the objective is to minimise supply shortfalls and obtain as much energy as possible. 10 reservoirs, 2401 variables, 1911 equality constraints and

4312 inequality constraints are involved. Penalty functions are used on soft constraints, Lagrange multipliers on hard constraints, while the elimination of variables by solving linear equations enforces some constraints. The resulting non-linear program is solved by Fletcher-Reeves conjugate directions.

The works of Dillon and Morsztyn (1972) and Dillon (1974) may not belong in this group of realistic system models, but they are based on one of their studies on the Tasmanian power system. It consists of seven hydro stations, and one thermal station. Actual thermal costs are replaced by a quadratic approximation. Two of the hydro stations are only run-of-river. The transmission system is modelled explicitly, including some transformers, but each busload is represented by a single smooth curve over the whole 12 months of the optimisation. So daily or weekly load peaks are not represented in any way. The time resolution is unclear, but appears to be a month. Pontryagin's Maximum Principle is used.

Hanscom *et al* (1980) modelled the Hydro-Quebec system of 7 reservoirs with 52 weekly time steps for the one year medium term problem. A stochastic dynamic program with fewer reservoirs, and a one month time step over a 7 year horizon provides end of year water values for the medium term problem. Results from this problem are themselves used as constraints on an hourly scheduling method with a one week horizon (the short term problem). Only the first week's results of the medium term problem are utilized, the algorithm then being resolved. A reduced gradient solution method is used. The cost function is a piecewise quadratic. This algorithm is a part of the Hydro-Quebec planning information system and a 0.1% saving in costs is estimated as being required to cover its development costs.

Rosenthal (1981) developed a non-linear network flow algorithm, solved by reduced gradients. It has been tested on a 6 reservoir section of the 19 reservoir Tennessee Valley Authority system.

These last four methods have tackled realistically large and difficult problems, and so the methods used are more likely to be generally applicable. However none appear to be actually used for power system control purposes.

2.4 STOCHASTIC METHODS

Six methods are briefly described which explicitly take account of the uncertain nature of water inflows. A seventh method is described by Quintana *et al* (1979), Chikhani *et al* (1979), and Quintana *et al* (1981) but none of these three papers elaborate sufficiently on the solution method (as opposed to modelling aspects) to be able to assess its usefulness or validity.

Stochastic dynamic programming is an obvious choice, but is limited in the dimension (number of reservoirs) that can be handled. Arvanitis and Rosing (1970) aggregate the U.S. Pacific Northwest system into a single reservoir, while Viramontes and Hamilton (1978) work with an arbitrary single reservoir model. This might be satisfactory for a system with high spatial correlation of water inflows permitting easy disaggregation of the system to get individual reservoir schedules. Its usefulness is doubtful when different reservoirs have quite different inflow patterns.

The linear decision rule described by Revelle *et al* (1969), Revelle and Gundelach (1975), is used in various forms by many authors interested in water resources management. With this scheme water releases are linearly related to reservoir storage. Revelle's principal concern is determining optimal reservoir size, assuming a good management strategy. The linear decision rule is not claimed to be the optimal rule, but is convenient to work with. Two methods of solution are evaluated - the use of deterministic optimisation with 20 years of streamflows, and chance constrained programming.

Electricité de France (Read 1979, Daellenbach and Read 1976) used decomposition by prices as mentioned in section 2.2. Two modifications were considered to take account of inflow variability. The first involved a number of deterministic optimisations for different historical inflow data years. The optimal solution was taken as the average of these. Each separate optimisation is carried out on the basis of perfect future knowledge. So the overall solution will be less cautious than in a situation where only imperfect inflow information is available. The second modification also used a number of years' inflow data records. Separate prices were specified for each. The individual reservoir subproblems each determine one storage trajectory. Different releases are then found for each inflow sequence which adhere to this trajectory as closely as possible. So thermal generation is now required to provide various outputs

depending on inflow levels, to compensate for the different releases with different inflows. Obviously reservoir levels should vary, to some extent, with inflow levels, and this method in contrast to the first gives excessively costly solutions.

Peters *et al* (1978) formulated a chance constrained non-linear program with recourse actions. Their water resource system involved three reservoirs in Iran, and was optimised over 12 intervals. The end of month reservoir target volumes are linearly dependent on inflows for the current month (the recourse action) and the previous two, as inflows are highly correlated over three months. The system here is quite unusual, so the method would not have wide application. Inflows into the New Zealand system are not highly correlated over such long periods, as with the Iranian model - such good three month correlations seem exceptional. Also, recourse actions could be difficult for hydro reservoirs with one week time steps due to the difficulty in making sufficiently accurate flow measurements.

Dillon *et al* (1980) optimise a linear model of a hydro-thermal system with a one month time step. Chance constrained and two stage linear programming with recourse methods were used. A special cost is applied to reservoir spills to ensure they are minimised, and losses are not handled. Again the difficulty in making measurements for recourse action decisions exists, and the model is restricted to linear features only.

2.5 A DECOMPOSITION METHOD

Decomposition was considered as a possible solution method before that of Chapter 3 was found to be more practicable. The solution structure of a decomposed method consists of a number of subproblems to be solved independently, with a co-ordinator examining their solutions. On the basis of overall system objectives, the co-ordinator adjusts some parameters which are then used by the sub-problems to obtain a new set of solutions. This process continues until co-ordination is satisfactory.

In our case, the problem can be formed by treating each hydro reservoir as a sub-system, with the co-ordinator ensuring that total generation equals demand. All the thermal stations combined form a subsystem. The problem we wish to solve is:

$$\min \sum_{t=1}^{13} F(\text{Thermal Generation } (t)) \tag{2.1}$$

subject to:

Total Generation (t) $\geq$ Demand (t) and various constraints on generating capacities, water storages etc., at individual reservoirs. (2.2)

Adjoining eqn (2.2) to eqn (2.1) by Lagrange multipliers:

$$L(G_{Th}(t),\lambda(t)) = \sum_{t=1}^{13} \{F(G_{Th}(t)) + \lambda(t)\left[D(t) - \sum_{v=1}^{m} G_v(t) - G_{Th}(t)\right]\} \tag{2.3}$$

with constraints on individual reservoirs and stations

where: $G_{Th}(t)$ = total thermal generation

$G_v(t)$ = hydro generation from reservoir v.

$\lambda(t)$ = Lagrange multiplier

$D(t)$ = demand

$F(G_{Th}(t))$ = cost of generating $G_{Th}(t)$

A saddle point $(G^*_{Th}(t), \lambda^*(t))$ is defined as:

$$\begin{aligned} L(G^*_{Th}(t),\lambda^*(t)) &\leq L(G_{Th}(t),\lambda^*(t)) \\ L(G^*_{Th}(t),\lambda^*(t)) &\geq L(G^*_{Th}(t),\lambda(t)) \end{aligned} \tag{2.4}$$

i.e. given the optimal $\lambda(t)$, $G^*_{Th}(t)$ is chosen to minimise L, or given the optimal $G_{Th}(t), \lambda^*(t)$ is chosen to maximise L. This suggests the solution method:

(i) Choose the initial values $\lambda^o(t)$, all $\lambda^o(t) > 0$. Set $j = 0$.

(ii) Solve the Lagrangian, eqn (2.3), with $\lambda(t) = \lambda^j(t)$ obtaining $G^j_v(t)$ for $v = 1,\ldots,n$ and $G^j_{Th}(t)$.

(iii) Calculate

$$h(\lambda^j(t)) = \sum_{t=1}^{13} \{F(G_{Th}(\lambda^j(t)) + \lambda^j(t)\left[D(t) - \sum_{v=1}^{m} G^j_v(\lambda^j(t)) - G^j_{Th}(\lambda^j(t))\right]\} \tag{2.5}$$

(iv) Find $\lambda^{j+1}(t) = \lambda^j(t) + \alpha^j$ where α^j is selected to maximise $h(\lambda^j(t))$ by minimising the mismatch between demand and total generation.

(v) Stop if mismatch is sufficiently small, otherwise return to (ii)

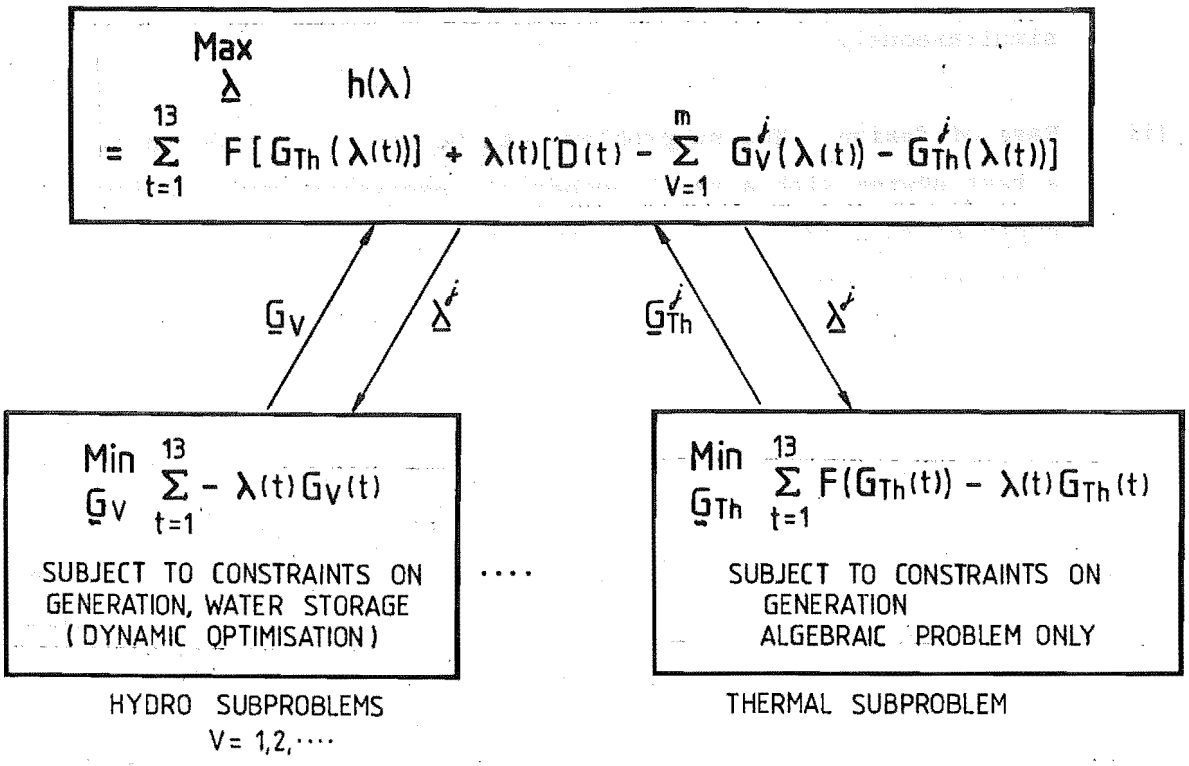


FIGURE 2.1 Structure of the Decomposed Problem

Step (ii) can be performed as a set of subproblems, as in Figure 2.1, as the Lagrangian is separable. The multipliers, $\lambda(t)$, can be interpreted as incremental costs of generation, or "price". The thermal subproblem is simple - generation with incremental cost less than or equal to $\lambda(t)$ is operated. Hydro generation is more difficult. These subproblems, finding the optimal $G_v(t)$ sequence for the given $\lambda(t)$ set handed down by the co-ordinator, make up the bulk of the work in finding the solution.

To summarise, the subproblems are solved with a given set of incremental costs, producing generation schedules. The co-ordinator looks at the total generation for each time interval, then adjusts $\lambda(t)$ upwards to increase generation if the total is less than demand, or vice versa. When demand is satisfied closely enough, the process stops.

Advantages offered by this approach include:

(i) Possibility of parallel processing. All subproblems could be solved simultaneously.

(ii) Ease of design. The subproblems can be developed independently. A test system with a small number of reservoirs could easily be expanded to handle whatever number required, or changed as the power system expands.

(iii) Economy of computer processing time. Fletcher and Reeves (1964) claim that their conjugate gradients method is guaranteed, apart from rounding errors, to locate the minimum of any quadratic function of N arguments in at most N interations. Hence a system of 8 subproblems, each of dimension N, might be expected to require computer time proportional to $(8N)^2$ if solved as a whole by this method. If decomposed, the 8 individual subproblems might on a similar basis require process time proportional to $8\ N^2$, for each price interation. If less than 8 price iterations are required, decomposition offers a saving. It will be shown that this anticipated economy in processing time could not be realised with this problem.

The hydro subproblems, as posed, are almost linear. If spilling is not permitted, or does not occur, they are linear. As a result, if $\lambda(t_1) = \lambda(t_2)$ then $G_v(t_1)$ and $G_v(t_2)$ might be able to take on a whole range of values giving the same subproblem cost, i.e. increasing $G_v(t_1)$ while decreasing $G_v(t_2)$ by the same amount will have no effect on subproblem v cost, but will affect $h(\underline{\lambda}(t))$, the co-ordination problem. This is because the demand/generation mismatch will be affected.

Figure 2.2 shows this difficulty for a simple two time step subproblem. The cost to be minimised is:

$$F_v = \lambda(1)G_v(1) + \lambda_2 G_v(2) \qquad (2.6)$$

where $x_v(0)$, $x_v(2)$ are given

$$x_v(1) = x_v(0) + i_v(1) - G_v(1)$$

$$x_v(2) = x_v(1) + i_v(2) - G_v(2)$$

$$0 \leq G_v(y) \leq \bar{G}_v$$

x_v = storage

i_v = inflow

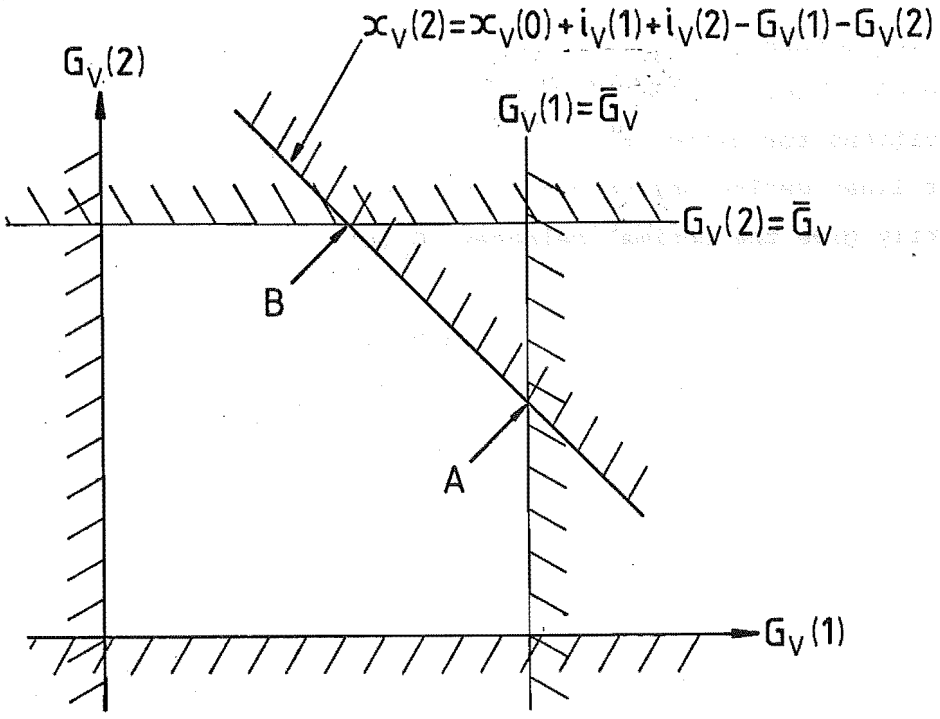


FIGURE 2.2 Two time step linear subproblem showing degeneracy difficulty.

The infeasible regions are shaded out in Figure 2.2. If: $\lambda(1) = \lambda(2)$ the solution $(G_v(1), G_v(2))$ lies anywhere along the line AB, i.e. a degenerate solution.

$\lambda(1) = \lambda(2) + \varepsilon$ => solution is at A.

$\lambda(1) = \lambda(2) - \varepsilon$ => solution is at B.

Hence the price adjustment mechanism cannot be sure of reaching a solution as the solution might be a specific point between A and B. Introducing a non-linearity, such as quadratic transmission losses, would put a curve into the cost function, eqn (2.6). As this is only a small effect, the curvature would not be great, and convergence is likely to be slow or unreliable. Losses also result in (relatively weak) coupling between subsystems, with non-zero off diagonal terms in $\frac{\partial^2 F_V}{\partial^2 x_i x_j}$.

From Lasdon (1970), the method will, in general, converge if $h(\underline{\lambda}^*(t))$ is differentiable, where $\underline{\lambda}^*(t)$ is the value of $\underline{\lambda}(t)$ at the solution. One of the conditions for differentiability is that $F(G_{Th}(t))$ be strictly convex. Straight lines define only a convex region, so the optimal $\lambda^*(t)$ does not necessarily give the optimal releases, $G_v(t)$.

Lasdon also points out that linear subsystems can only arrive at activity levels which are extreme points of their constraint sets. The optimal activity levels of the overall system may be interior to the sub-system constraints and only by weighting the various extreme points can these interior points be generated.

So complete decomposition of a linear system is not possible. Dantzig-Wolfe solution by decomposition of linear programs is actually only partial decomposition, functioning in the manner outlined by Lasdon, above, i.e. the co-ordination level weights subsystem solutions.

As the problem is almost linear, it was found that the expected square relationship between problem dimension and computer time does not apply. The model described in Chapter 3 has non-linearities due to the spilling of excess water, and quadratic transmission line losses. Three and eight reservoir models have been solved (neither using decomposition), by the same conjugate gradients method. The eight reservoir problem required about 15% more function evaluations than the three reservoir version. Hence processing time is approximately proportional to problem dimension, and the square relationship mentioned earlier does not apply. This indicates that each price interation of a decomposed problem would require as much computer time as the complete solution without decomposition.

Difficulties caused by the "nearly linear" nature of the problem led to the rejection of the decomposition approach. The lack of any savings in computer time did not become evident until much later. Modelling transmission losses produces some curvature in the sub-problem objective functions, but this is only a small effect. It appeared to be insufficient to prevent small changes in price from one iteration to the next causing switching from one corner point to another (as for the completely linear case). As a result, attention turned to the method of Chapter 3, solving one large non-linear program by a reliable conjugate gradients algorithm, described in Chapter 4.

2.6 CONCLUSIONS

The methods reviewed generally fell into four groups:

(i) Deterministic, excessively simple problems, so the method is unproved or inapplicable for large, realistically complicated systems.

(ii) Deterministic, realistic models, with the potential to assist power system operating organisations to varying extents.

(iii) Stochastic one reservoir models giving a true stochastic solution for that model, but not applicable to multi-reservoir systems.

(iv) Multi-reservoir stochastic methods neglecting various aspects such as transmission losses, or some non-linear effects, or only crudely approximating stochastic inflows.

Experience with the decomposition method of section 2.5 led to the choice of a single large nonlinear programming solution technique as most likely to succeed. Also, the non-linear programming approach is the least restrictive on the model, of the methods considered. Decomposition was considered likely to be unreliable, have greater difficulty in handling coupling between reservoirs (non-separable benefits), and to have, at best, advantages that are declining in significance as computers become more powerful.

CHAPTER 3

OUTLINE OF A GENERAL MODELLING AND SOLUTION TECHNIQUE FOR RESERVOIR SCHEDULING

3.1 INTRODUCTION

In Chapter 2 various possible approaches to hydro-thermal scheduling were discussed. It was concluded that a mathematical programming method was most likely to succeed. A generalised description of the deterministic hydro thermal scheduling model will be given in this chapter. Chapter 5 presents the application of these principles to the New Zealand power system, and includes extensions to handle a D.C. link.

The algorithm is designed for use over a one year time horizon, with a time step of from one to four weeks.

Modelling of physical components of the power system will be described first, including details of the development of the D.C. loadflow loss coefficients. The means of converting a general time dependent problem into a static equivalent by forming a Hamiltonian follows. Then this method is applied to the model, constraint enforcement described, and gradient calculations given. The result is a set of equations suitable for solution by an unconstrained static hill climbing algorithm. Chapter 4 describes the conjugate gradients method used to do this.

Throughout the model development the most simplifying assumptions possible have been made, unless there is a good reason for doing otherwise.

Only those components of the power system likely to be limiting factors need be modelled. If some factor is not thought to be worth considering by those involved in the present planning methods, then it is likely not to be worth including in the model. If well designed,the power system will not be limited in its performance by, for example, the rating of a single transformer at some substation, so there is no need for it to be modelled.

The level of model detail must be consistent with the time scale of the problem. It is not possible to ensure that all components of the system will always be within their normal operating range when dealing with a one week time step as the exact load pattern is not known, nor could it be modelled. Also, there is no point in devising a model so complicated that it can not be solved.

Consequently transmission system constraints have not been modelled, but constraints on power transfer between two regions can be incorporated, as is done for the New Zealand example. Transmission losses are modelled by the D.C. loadflow loss coefficients calculated before the optimisation commences. Losses are thereafter calculated from the generator outputs and these coefficients only.

All thermal stations and large reservoirs are modelled individually. Reservoirs in series are possible, and uncontrollable tributary inflows can be handled. Run-of-river hydro stations are modelled with little effort. Constraints on reservoir levels, reservoir releases, and generation can be enforced, and may be time varying.

3.2 DEMAND

Power demand and other time-dependent quantities are expressed as average values over discrete time intervals. One week is likely to be the most convenient interval in practice, as loads have a weekly cycle. Short duration peaks often require the use of higher cost thermal stations than consideration of average weekly energy requirements would suggest. Some means of representing load fluctuation during each week is needed.

Figure 3.1 is a load duration curve, giving the demand $A_t(h)$ in MW which is exceeded for h hours, in week t. This curve is approximated by ℓ load classes. (ℓ=3 for example of Chapter 5). The load classes can, and will usually, be selected to be of different durations. The peak load segment might be made shorter than the others to ensure that the weekly maxima are represented well, as these can require very expensive generation.

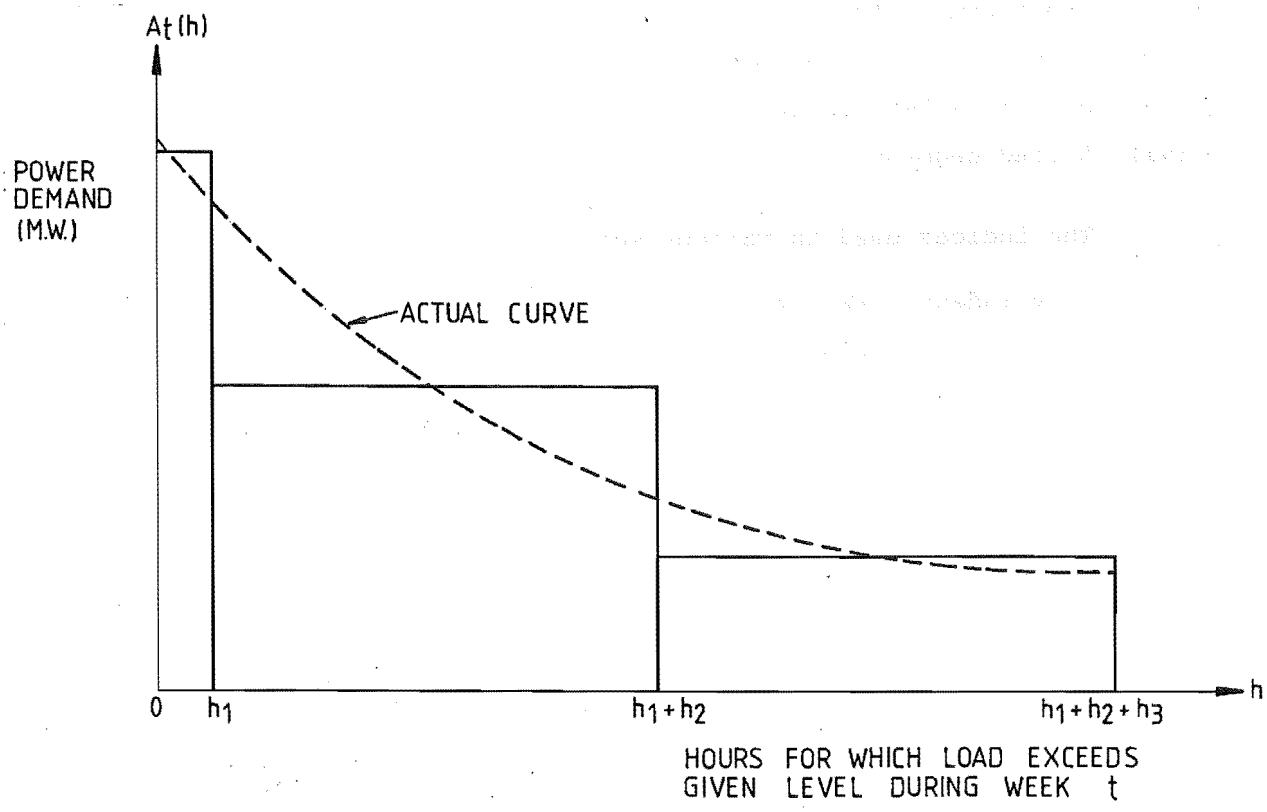


FIGURE 3.1 Load duration curve and three segment approximation

The MWh, or energy, consumed during each load class is the area under the curve for the appropriate time, e.g.

$$D(t,j) = \int_{h_1+\ldots+h_{j-1}}^{h_1+\ldots+h_j} A_t(h)\,dh \qquad (3.1)$$

where $D(t,j)$ is the demand for time interval t, load segment j.

h_j = duration of the j^{th} load segment.

3.3 GENERATION AND STORAGE

3.3.1 Introduction

The model is formulated for a system consisting of:

(i) m reservoirs, m-2 of which are hydraulically independent.

(ii) reservoir q discharges into reservoir q+1 (where $q+1 \leq m$). This relationship is chosen to apply to reservoirs q and q+1 only for notational convenience.

(iii) n-m thermal stations.

(iv) p run of river stations.

(v) 52 time intervals (weeks).

(vi) ℓ load segments.

The indices used in various summations include:

v indexing reservoirs and thermal stations,

t indexing time segments,

j indexing load segments.

Specific features of each type of station follow.

3.3.2 Hydro Reservoirs

The most important assumption made in this model is that all hydro stations fed by a given long term reservoir, but upstream from the next long term reservoir if it exists, can be lumped into one equivalent station. It is assumed that the individual scheduling of stations, from an aggregated figure, can be done without upsetting the long term schedule significantly. This process would require another optimisation with a shorter time scale.

Generation at a hydro station is a function of head, flow rate, and tail water elevation. Figures 3.2 and 3.3 show how the water to electrical power conversion factor (in CUSECS/MW) varies with generation for Benmore and Waitaki stations. The various curves are for when all machines actually in use are equally loaded, as this gives the most efficient operation by the equal incremental cost principle.

Clearly these characteristics are impossible to model for weekly outputs as the exact loadings on each station are not known - only the average output of the whole chain of stations over each load segment. Even when a reservoir feeds only one station, the fact that average, not instantaneous, loadings are calculated by the scheduling algorithm prevents the use of exact conversion factors.

Head variations are only known for stations drawing water directly from their long term storage reservoir. At other stations head effects can not be modelled, as the variations are unknown.

Using the available information, some approximations must be made. Here, head effects have been neglected, and constant water to electrical energy conversion factors have been used. This permits considerable simplification. So:

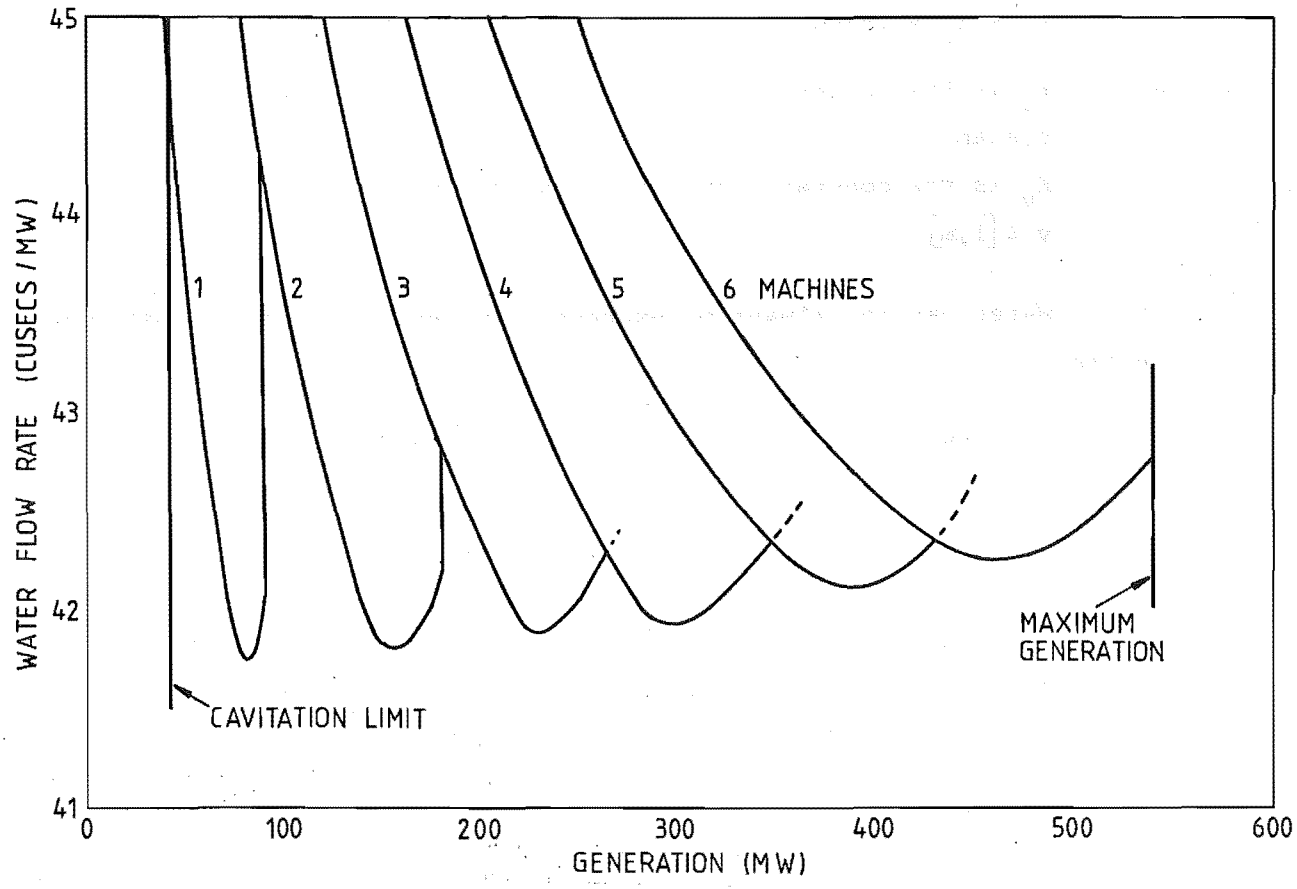


FIGURE 3.2 Conversion efficiency, water to electricity, at Benmore (Green, 1971b).

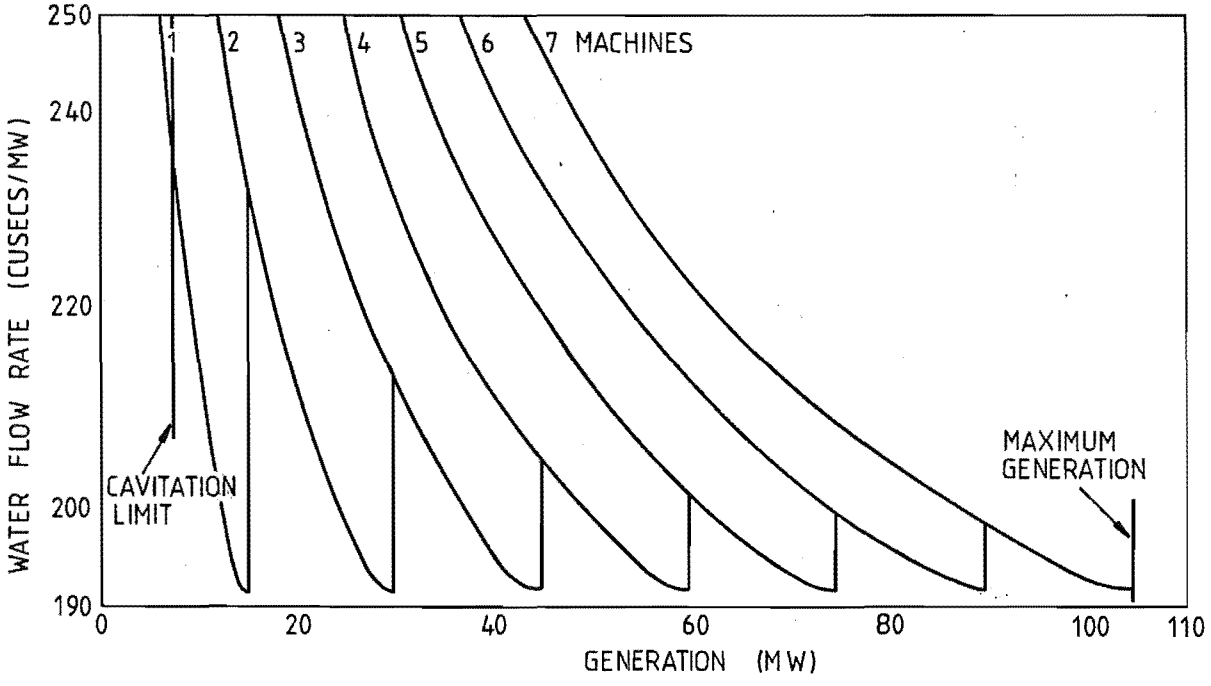


FIGURE 3.3 Conversion efficiency, water to electricity, at Waitaki (Green 1971b).

$$R_v = K_v \times \text{volume of water released} \tag{3.2}$$

where: R_v is the energy generated, in MWh, by all stations in a river system,

K_v is the constant conversion factor,

$v \in [1,m]$

Water can now always be expressed in terms of its generation potential in MWh.

A further simplification is that leakage and evaporation losses from large storage reservoirs need not be considered in most cases. This is because lake inflows are measured by recording lake level and outflow. No other method is practical for a large lake with many small rivers flowing into it. Hence the lake's losses do not appear at all, being automatically subtracted from inflows. This will be an accurate result if the surface area of the lake does not change too much with its level.

The water storage equation for all lakes except q+1, with all quantities expressed as their electrical equivalents in MWh is:

$$X_v(t+1) = X_v(t) - \sum_{j=1}^{\ell} R_v(t,j) + I_v(t) - S_v(t) \tag{3.3}$$

and

$$G_v(t,j) = R_v(t,j) + T_v(t) \frac{h_j}{\sum_{i=1}^{\ell} h_i} \tag{3.4}$$

where $X_v(t)$ = storage in lake v at beginning of time interval t.

$R_v(t,j)$ = release from lake during time t for load segment j.

$I_v(t)$ = storable inflow.

$S_v(t)$ = spill.

$G_v(t,j)$ = generation from all stations fed by the reservoir.

$T_v(t)$ = tributary flows, non storable, but may be non-existant for some reservoirs.

h_j = duration of load segment j (hours).

Spill is calculated from

$$S_v(t) = \begin{cases} X_v(t) - \sum_{j=1}^{\ell} R_v(t,j) + I_v(t) - \bar{X}(t) \\ \qquad \text{if } X_v(t) - \sum_{j=1}^{\ell} R_v(t,j) + I_v(t) > \bar{X}_v(t) \\ 0 \text{ otherwise} \end{cases} \tag{3.5}$$

i.e. lake levels are simply clipped off at their maximum values, the excess being "spill". This strategy is in contrast to the approach of, for example, Dillon (1974) who applies a penalty to minimise spill. Our strategy is justified in section 3.6.5.

Constraints on the variables of eqns (3.3) and (3.4) must be enforced. Lower limits on storage exist.

$$\underline{X}_v(t) \leq X_v(t) \tag{3.6}$$

where $\underline{X}_v(t)$ = minimum storage, which may be time varying.
Limits on generation are:

$$\underline{G}_v(t,j) \leq G_v(t,j) \leq \bar{G}_v(t,j) \tag{3.7}$$

where $\underline{G}_v(t,j)$ = minimum generation,

$\bar{G}_v(t,j)$ = maximum generation,

and for reservoir releases:

$$\underline{R}_v(t,j) \leq R_v(t,j) \leq \bar{R}_v(t,j) \tag{3.8}$$

where $\underline{R}_v(t,j)$ = minimum release,

$\bar{R}_v(t,j)$ = maximum release.

The releases, $R_v(t,j)$, are the control variables in this model, so restrictions on generations, $G_v(t,j)$, must be converted into constraints on $R_v(t,j)$ for the optimisation procedure. The state (dependent) variables are reservoir storages.

Tributary flows are assumed to occur uniformly over the whole of each time interval. The limited storage behind individual dams is assumed to be sufficient to smooth out fluctuations over a time interval. Spilling of tributary flows is allowed only when they are so high that generation limits are exceeded with minimum permitted controllable releases, i.e.

$$T_v(t) = \bar{G}_v(t,j) - \underline{R}_v(t,j)$$
$$\text{if } T_v(t) \frac{h_j}{\sum_{i=1}^{\ell} h_i} > \bar{G}_v(t,j) - \underline{R}_v(t,j) \text{ for any } j. \tag{3.9}$$

Tributary flows are considered to be fully utilised otherwise, imposing a restriction on the minimum possible generation:

$$\underline{G}_v(t,j) \geq T_v(t) \frac{h_j}{\sum_{i=1}^{\ell} h_i} + \underline{R}_v(t,j) \qquad (3.10)$$

Maximum reservoir releases are also restricted - they can not exceed generation capacity remaining after tributary flows have been utilised:

$$\bar{R}_v(t,j) \leq \bar{G}_v(t,j) - T_v(t) \frac{h_j}{\sum_{i=1}^{\ell} h_i} \qquad (3.11)$$

For the lower reservoir of the cascaded pair of reservoirs only equation (3.3) need be changed:

$$X_{q+1}(t+1) = X_{q+1}(t) - \sum_{j=1}^{\ell} R_{q+1}(t,j) + I_{q+1}(t) - S_{q+1}(t)$$

$$+ \sum_{j=1}^{\ell} G_q(t,j) \cdot k_{q/q+1} \qquad (3.12)$$

where $k_{q/q+1}$ converts generation from stations fed by reservoir q to give the generation potential of the water they have used when it flows into reservoir q+1.

Equation (3.12) assumes that the spill $S_q(t)$ is lost completely, but the equation could be modified so that spill also goes into reservoir q+1. The time delay of flow from one reservoir to the next is neglected - valid if it is much less than a week.

3.3.3 Run-of-River Hydro

Any hydro station with only short term storage facilities, or any other station for which output can be predetermined is treated in a very simple way. This category can include for example geothermal stations, which have zero marginal cost, and so will be operated whenever available.

The output of these stations is treated as negative demand. Total run of river output is therefore subtracted from total demand figures for each time and load segment. When calculating loss coefficients, their outputs appear as negative bus loads.

3.3.4 Thermal Power Stations and Their Cost

The thermal station model is very simple. Equation (3.13) gives thermal generation requirements

$$\text{Thermal Generation } (t,j) = D(t,j) - \sum_{v=1}^{m} G_v(t,j) - \sum_{RR=1}^{p} G_{RR}(t,j) + L(t,j) \qquad (3.13)$$

where $D(t,j)$ = demand for time t, load segment j

$L(t,j)$ = transmission system loss,

G_v = controllable reservoir generations,

G_{RR} = run-of-river generations.

Thermal stations are loaded as efficiently as possible until the required output is reached. A fictitious station of extra high cost is added to account for any shortfall in supply.

As for the hydro stations, exact loadings are not known, so some approximations must be made in costs. For example, should thermal units reach maximum efficiency at full load, then the assumption can be made that all machines are run either at full load, or not at all. This conveniently gives a constant incremental fuel cost, so

Total Cost of Thermal Generation for (t,j)

$$= \sum_{v=m+1}^{n+1} \lambda_v G_v(t,j) \qquad (3.14)$$

where λ_v is the incremental cost of station v in \$/MWh.

$G_{n+1}(t,j)$ is the shortfall in generation. This cost can also be expressed in terms of the optimisation's control variables (reservoir releases):

$$F\Big[\sum_{v=m+1}^{n+1} R_v(t,j)\Big]$$

$$= F\Big[D(t,j) - \sum_{v=1}^{m} \Big\{R_v(t,j) + T_v(t) \frac{h_j}{\sum_{i=1}^{\ell} h_i}\Big\} - \sum_{RR=1}^{p} G_{RR}(t,j) + L(t,j)\Big] \qquad (3.15)$$

which cost, when summed over the year, is the quantity to be minimised by the optimisation, subject to constraints. The function F[ ] gives the total cost of the specified thermal generation, assuming the stations are loaded as efficiently as possible. It is a piecewise linear function, in thermal generation, if eqn (3.14) holds, but there is no reason why any appropriate increasing function can not be used. All that is needed is a specified loading order for thermal units to determine this function.

3.3.5 Equipment Outages

Generation equipment outages are treated in a similar way for all types of stations. Planned maintenance downtimes, known in advance, can be represented by reducing generation capacity at the appropriate times. Forced outages can be estimated, perhaps using historical data, and installed capacity reduced by the average value.

For some hydro subsystems, discharge constraints from the long term reservoir may be the limiting factor on generation, rather than installed capacity. This can conveniently eliminate the need to consider outages, unless they are very serious.

3.4 TRANSMISSION LOSS BY D.C. LOADFLOW LOSS COEFFICIENTS

The object of loss coefficient methods is to give a simple formula involving only generator outputs for determining total system transmission losses. They also enable the calculation of the change in losses with respect to changes in generation. Effectively the system's physical structure and the loads at each bus are replaced by a set of coefficients, calculated infrequently but used often.

The usual method of determining transmission losses for power system optimisation problems is by using the B-coefficients (Kirchmayer, 1958). This permits losses to be expressed as a function of generator outputs:

$$L = \underset{\sim}{G}^t \, B_{vv} \underset{\sim}{G} + \underset{\sim}{G}^t \, \underset{\sim}{B}_{v0} + B_{00} \qquad (3.16)$$

where L = transmission loss.

$\underset{\sim}{G}$ = vector of generator powers.

B_{vv}, $\underset{\sim}{B}_{v0}$ B_{00} a matrix, a vector, and a scalar respectively are the B-coefficients.

The superscript t denotes vector transpose.

Computation of these coefficients is not straightforward - data from A.C. loadflows is needed - and some unrealistic assumptions must be made about the system:

(i) The equivalent load current at any bus remains a constant complex fraction of the total equivalent load current.

(ii) Generator bus voltage magnitudes remain constant.

(iii) The ratio of reactive power to real power of any source remains a fixed value.

(iv) Generator bus angles remain constant.

As the optimisation proceeds from an initial guess for reservoir releases to a final solution, these conditions will be violated. Hence frequent recalculation of coefficients is necessary if they are to remain valid.

Podmore (1972, 1973) developed a considerably simpler method based on the D.C. loadflow equation. Calculation of these loss coefficients requires only line resistance and bus load power data. Supplementary data from A.C. loadflows is not necessary. As the coefficients are calculated from constant data, updating as the optimisation proceeds is not necessary. One set of loss coefficients is required, however, for each time interval and load segment, and also a different set is needed for each possible slack bus.

Total system losses are calculated from:

$$L = \underset{\sim}{G}^{t}\ ECC\ \underset{\sim}{G} + G^{t}\ E\underset{\sim}{C}0 + E00 \qquad (3.17)$$

where ECC is a square matrix of loss coefficients, independent of load powers.

$E\underset{\sim}{C}$ and $E\underset{\sim}{C}0$, a vector and a scalar respectively are load dependent coefficients. The slack bus power, being dependent on total losses, does not enter into the calculation of equation (3.17). ECC must be calculated for each possible slack bus, but $E\underset{\sim}{C}0$ and E00 must be calculated for each slack bus and also each time interval and load segment.

The coefficients are derived as follows:

(i) System losses are separated into voltage magnitude and angle dependent components:

$$L = \underline{\delta}^t C_{NN} \underline{\delta} + P_E \tag{3.18}$$

where $\underline{\delta}$ is a vector of bus angles

P_E is the voltage magnitude dependent loss

C_{NN} is a matrix of branch conductances

(ii) Bus voltage magnitudes are assumed to be constant and only angle dependent losses are considered.

(iii) The D.C. loadflow equation is applied to express the bus angles linearly in terms of the bus powers.

$$\underline{\delta} = Z_{NN} \underline{G}_N \tag{3.19}$$

By substituting in eqn. 3.17 for $\underline{\delta}$

$$L = \underline{G}^t Z_{NN} C_{NN} Z_{NN} \underline{G} + P_E \tag{3.20}$$

where $Z_{NN} = B_{NN}^{-1}$

and B_{NN} is a matrix of branch susceptances

$\underline{G}$ is a vector of bus loads and generator power inputs

(iv) Separate load and generator powers, substitute for the loads, (which we have data for) to obtain:

$$L = \underline{G}^t ECC \underline{G} + G^t EC\underline{0} + E00 + P_E \tag{3.21}$$

Making an approximation by setting bus voltages to 1 p.u. gives $P_E=0$, and hence eqn. (3.17), which expresses losses in terms of generator powers. If it is desired to treat a load explicitly, it can be regarded as a generator with a negative output. Hence it will appear explicitly in eqn. (3.17) rather than in the derivation of the coefficients.

Likewise, run-of-river hydro stations and other generating sources, the output of which is to be specified in advance rather than optimised, can be treated as negative loads. They are then absorbed into the loss coefficients, reducing the number of coefficients to be stored.

TABLE 3.1 Tests on Loss Coefficients

System	No. of Buses	No. of Generators	Typical X/R	Exact Losses	Loss Coefficient Losses	% Error
A.G.E.	28	8	3	41.59	40.57	2.4
N.Z.E. (220 kV)	27	7	5	42.78	44.63	4.3
N.Z.E. (South Island)	185	8	5,3,2	83.28	87.71	5.3
39 Bus	39	10	10	34.32	34.97	1.9
I.E.E.E.	118	18	3	132.65	126.11	4.9

Table 3.1 shows some tests on various systems performed by Podmore (1972). It can be seen how the best results are obtained for a system with high X/R ratio, i.e. a modern high voltage system. The 5.3% error for the N.Z.E. South Island system is probably similar to that for the North Island, which seems a reasonable accuracy for the optimisation problem.

This method has another important feature making it especially suitable for the hydro thermal algorithm - the slack bus power is the dependent variable. On the other hand, the B-coefficient method has the total system load as dependent variable. Now, the highest incremental cost thermal generation is at the slack bus (any hydro station will do if no thermal generation applies). This is appropriate because this is the bus which takes up any changes in losses or hydro generation.

Therefore $\partial L/\partial G_v$ gives the incremental loss when an increment of power is sent from generator v to the slack generator. This is exactly what is required by the optimisation process gradient calculations. A reduction in output from one hydro station, keeping all others constant, causes an increase in output from the most expensive thermal station, due to the way in which thermal station outputs are calculated. The B-coefficients method gives $\partial L/\partial G_v$, the incremental loss when an additional increment of power is delivered to the total load. This is inappropriate as generation is just being redistributed, as the solution proceeds, to meet a known load, not increased as the B-coefficients gradients indicate.

3.5 THE PRINCIPLE OF SOLUTION BY HAMILTONIAN

The conversion of a multi-stage (i.e. discrete time) problem into a form suitable for the optimal solution to be found by a static hill climbing method will be described. This is taken from Bryson and Ho (1975). Consider the system, also shown in Figure 3.4.

$$x(i+1) = h^i(x(i), u(i)), \; x(1) \text{ given}, i=1,\ldots,N \tag{3.22}$$

With a performance index to be minimised

$$J = \phi[x(N+1)] + \sum_{i=1}^{N} f^i[x(i),u(i)] \tag{3.23}$$

Adjoin the system equations (3.22) to J with a multiplier sequence $\rho(i)$:

$$\bar{J} = \phi[x(N+1)] + \sum_{i=1}^{N} [f^i(x(i),u(i)) + \rho^T(i+1)(h^i(x(i),u(i)) - x(i+1))] \tag{3.24}$$

Define the Hamiltonian, a scalar sequence H^i:

$$H^i = f^i(x(i),u(i)) + \rho^T(i+1)h^i(x(i),u(i)) \quad i=1,\ldots,N \tag{3.25}$$

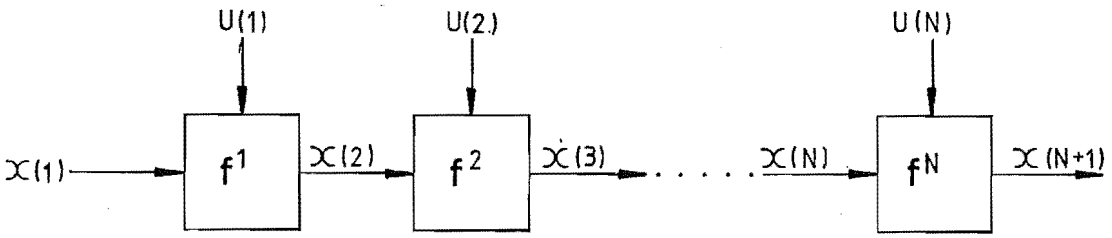


FIGURE 3.4 Multi-stage decision process u = control or decision, x = state.

$$\Rightarrow \bar{J} = \phi\left[x(N+1)\right] - \rho^T(N+1)x(N+1) + \sum_{i=2}^{N}\left[H^i - \rho^T(i)x(i)\right] + H^1 \qquad (3.26)$$

If the following choice for $\rho^T(i)$ is made, a useful result is obtained:

$$\rho^T(i) = \frac{\partial H^i}{\partial x(i)} = \frac{\partial f^i}{\partial x(i)} + \rho^T(i+1)\frac{\partial h^i}{\partial x(i)} \quad i=1,\ldots N$$

and $$\rho^T(N+1) = \frac{\partial \phi}{\partial x(N+1)} \qquad (3.27)$$

Now the cost function which is to be minimised has an extremum when

$$\frac{\partial H^i}{\partial u(i)} = 0 \qquad \text{for } i=1,\ldots,N \qquad (3.28)$$

So the optimal sequence, u(i), minimising the cost eqn (3.23), subject to eqn (3.22), is found by solving the two coupled difference equations (3.22) and (3.27) and looking for a point where (3.28) holds, with the boundary conditions for x(1) and ρ(N+1) given in (3.22) and (3.27).

The solution procedure for the hydro-thermal problem is shown in Figure 3.5. The state x(i) corresponds to reservoir storage levels. It is calculated in forward time, using equation (3.22), beginning with the given x(1). These results are used to determine the costate, ρ(i), in reverse time using (3.27), beginning with the given ρ(N+1). Now that all other conditions have been satisfied $\partial H^i/\partial u(i)$ can be calculated, and a new control strategy u(i) found in an attempt to get closer to the point where $\partial H^i/\partial u(i) = 0$. Controls, u(i), correspond to reservoir releases.

In the next section state and control variable limits are enforced. The cost function can then be given, and the equations for costate (ρ(i)) and gradient ($\partial H^i/\partial u(i)$), referred to above, can be found.

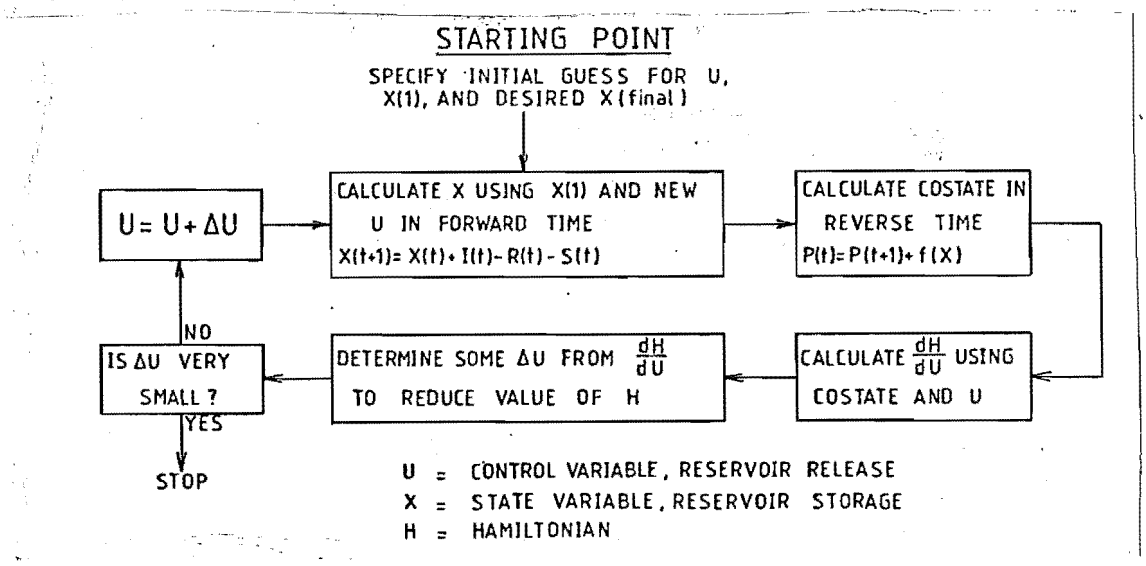


FIGURE 3.5 Solution procedure by Hamiltonian for the hydro-thermal problem

3.6 FORMULATION OF THE OPTIMISATION PROBLEM

3.6.1 State Constraints

State constraints consist of minimum reservoir levels, and the final time target level. The latter is only a minimum desired level, as it may be necessary to exceed it sometimes. Maximum levels are enforced by spilling. A penalty function is used for minimum state constraints, the following quantity being added to the cost function when a constraint is violated:

$$W_v(\underline{X}_v(t) - X_v(t))^2 \tag{3.29}$$

where W_v is the (scalar) weighting factor for reservoir v.

To avoid extremely steep contours, the weighting factors W_v are made small initially. A solution is obtained, and weighting factors increased (usually by a factor of 10) for those reservoirs not meeting constraints to the desired accuracy. The process then repeats. Penalty powers less than 2 were also tested, the results being given in section 6.5.3.

The terminal state target level is likewise enforced by

$$W_{(f)v}\ (\underline{X}_v(t_f) - X_v(t_f))^2 \tag{3.30}$$

where $\underline{X}_v(t_f)$ = final time target minimum

t_f = final time.

The use of separate weighting factors for minimum storages and final time target storage helps convergence as $(\underline{X}_v(t_f) - X_v(t_f))$ is likely to be much larger than any other value of $(\underline{X}_v(t) - X_v(t))$, to begin with.

3.6.2 Control Constraints

Two methods of control constraint enforcement were used - a transformation technique and penalty functions. The latter method uses equations similar to those for state constraints.

For violations of maximum release limits, the penalty is:

$$WU_v(\bar{R}_v(t,j) - R_v(t,j))^2 \tag{3.31}$$

where WU_v = weighting factor for reservoir v release violations.

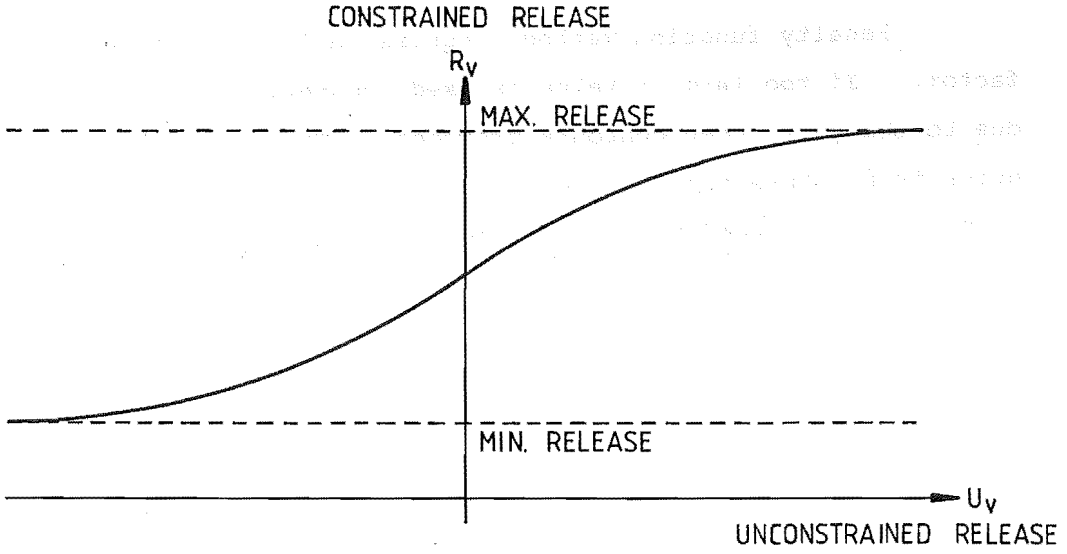


FIGURE 3.6 Assymptotic Transformation for controls.

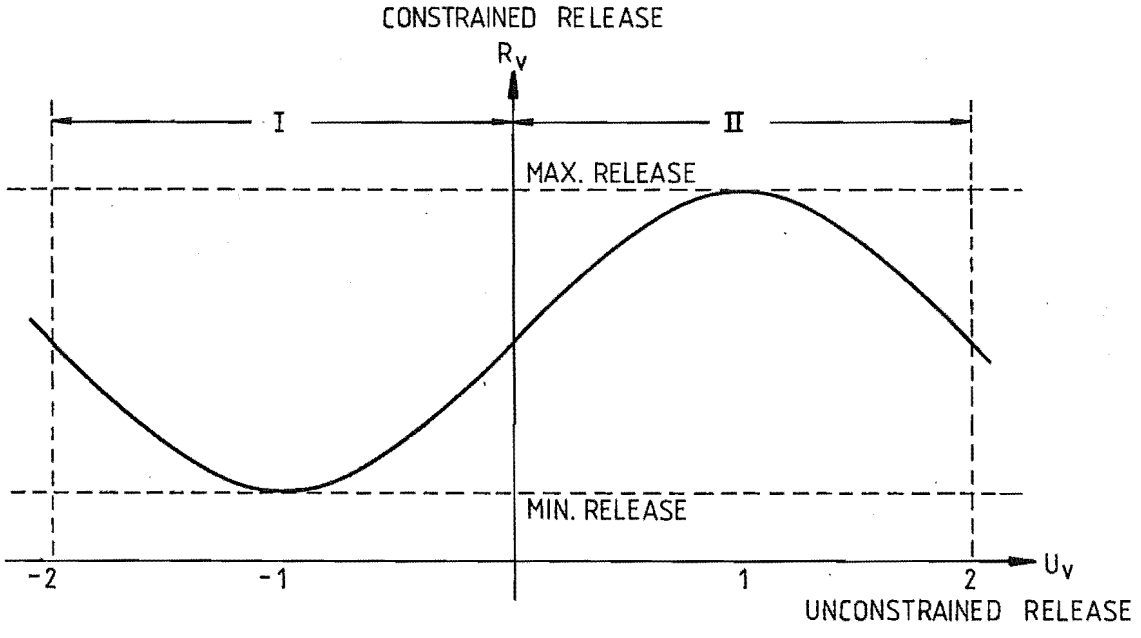


FIGURE 3.7 Piecewise parabolic Transformation for controls.

And for violations of minimum release limits:

$$WU_v(\underline{R}_v(t,j) - R_v(t,j))^2 \tag{3.32}$$

Penalty function methods require an initial value of the weighting factor. If too large a value is used, convergence may not be possible due to the very steep contours produced. To avoid difficulties that might arise from interaction between state and control penalties when constraints on both are enforced by penalty functions, a control transformation technique was tried first. (State constraints were still enforced by penalty functions). With this method the static optimisation algorithm performs searches in a space formed by an unconstrained control variable U. Before calculating state variables and determining the cost function, the constraining transformation is applied. This maps the unconstrained values into the constrained, physical, variables' permissible range.

At first a relationship forming asymptotes at upper and lower limits was tried (see Figure 3.6). This suffered from stalling when any particular release was required to move away from a large positive or negative unconstrained value, (i.e. very near a constraint). The slope of the transformation is very small here, so large changes in unconstrained values produce only a very small effect on the constrained equivalent.

The piecewise parabolic function of Figure 3.7 was finally adopted. This requires less computation than a trigonometric function which might seem a logical choice. The possibility of stalling on a constraint is easily overcome. A small perturbation is made to all releases that are close to a constraint at the beginning of each penalty function iteration.

Any variable which is within 5% of either of its extreme values is moved to 10% away from that extreme. If the solution really is so close to the extreme value, then the variable will simply move back again.

The transformations, UNSCALED, are:

$$\text{Region I:}\quad R_v(t,j) = \frac{\bar{R}_v(t,j)-\underline{R}_v(t,j)}{2}\,(1+2\,U_v(t,j) + U_v^2(t,j)) + \underline{R}_v(t,j) \tag{3.33}$$

$$\text{Region II:}\quad R_v(t,j) = \frac{\bar{R}_v(t,j)-\underline{R}_v(t,j)}{2}\,(1+2U_v(t,j)-U_v^2(t,j)) + \underline{R}_v(t,j) \tag{3.34}$$

where U_v = unconstrained control variable

Region I => $U_v \in [-2,0]$

Region II => $U_v \in [0, 2]$

Note that all unconstrained variables take values from the range -2. to 2. only. -1 gives minimum constrained release, and +1. the maximum. This results in a poorly scaled problem, i.e. a given change has much more effect on the objective function for some variables than the same change in others. Large reservoirs, releases from which affect the cost function greatly, then correspond to the steep sides of a valley, while small less significant reservoirs give a gently sloping valley floor.

This is the very type of problem troublesome for a steepest descent method, although it could be handled with ease by the conjugate gradients algorithm CGRADS (see Chapter 4). Steepest descent methods are likely to zigzag across the valley floor, and stop prematurely. The result is that large reservoirs would be scheduled efficiently, but small ones not. To enable comparisons, a square root scaling factor was applied to transformations, giving more nearly circular contours.

In equations (3.33) and (3.34) if U ranges from -1.0 to 1.0, then R_v goes from minimum to maximum release. For reservoir v, the gradient G_v is

$$G_v = f_e . f_v . \frac{\partial R_v}{\partial U_v} \qquad (3.35)$$

where f_e accounts for factors external to the reservoir (e.g. thermal fuel costs)

f_v accounts for reservoir constraints etc. and $\frac{\partial R_v}{\partial U_v}$ arises from the transformation equation relating R_v and U_v.

Then the increment in U_v over an iteration of the hillclimbing algorithm is

$$\delta_v = f_e . f_v . \frac{\partial R_v}{\partial U_v} . SZ \qquad (3.36)$$

where δ_v is the increment in unconstrained release

SZ is the step length, specified by the hill climbing routine.

If the percentage change in constrained releases for all the reservoirs is to be the same at each step, for say R_v half way between its limits (i.e. U_v=0)

$$\frac{\delta v}{D_v} = \text{constant for all } v \tag{3.37}$$

where D_v = range of the constrained release for reservoir v.

$$\Longrightarrow \frac{\delta R_v}{\delta U_v} / D_v = \text{constant (for } U_v = 0)$$

But for 3.33 and 3.34 this change is proportional to the range of permissible releases.

$$\frac{\delta R_v(t,j)}{\delta U_v} / D_v = \frac{\bar{R}(t,j) - \underline{R}(t,j)}{2} \tag{3.38}$$

However,

$$\frac{\delta R_v(t,j)}{\delta U_v} / D_v = \frac{2\sqrt{\bar{R}(t,j) - \underline{R}(t,j)}}{\sqrt{\bar{R}(t,j) - \underline{R}(t,j)}} \tag{3.39}$$

$$= 2$$

which is the desired property, holds for (3.40) and (3.41) below, the SCALED transformations:

Region I:

$$R_v(t,j) = \frac{MX_v(t,j)}{2}\left(1 + \frac{2\,U_v(t,j)}{\sqrt{MX_v(t,j)/2}} + \frac{U_v^2(t,j)}{(\sqrt{MX_v(t,j)/2})^2}\right) + \underline{R}_v(t,j) \tag{3.40}$$

Region II:

$$R_v(t,j) = \frac{MX_v(t,j)}{2}\left(1 + \frac{2U_v(t,j)}{\sqrt{MX_v(t,j)/2}} - \frac{U_v^2(t,j)}{(\sqrt{MX_v(t,j)/2})^2}\right) + \underline{R}_v(t,j) \tag{3.41}$$

where $MX_v(t,j) = \bar{R}_v(t,j) - \underline{R}_v(t,j)$

Region I => $U_v \in \left[-\sqrt{MX_v(t,c)},\ 0.\right]$

Region II => $U_v \in \left[0, \sqrt{MX_v(t,c)}\right]$

Now all reservoirs' releases should move from their initial guess values towards the solution at about the same rate, as the optimisation proceeds.

3.6.3 Summary of Model So Far

The cost of thermal generation is given by equation (3.15). To this must be added penalties for state constraint violations, equations (3.29) and (3.30). Constrained releases are now a function of their corresponding unconstrained variables, equations (3.40) and (3.41). The resulting cost function to be minimised is:

$$J = \sum_{t=1}^{52} \left\{ \sum_{j=1}^{\ell} \left[F\left[D(t,j) - \sum_{v=1}^{m} \left(R_v(t,j) + T_v(t) \frac{h_j}{\sum_{i=1}^{\ell} h_i} \right) \right. \right. \right.$$
$$\left. \left. \left. - \sum_{RR=1}^{P} G_{RR}(t,j) + L(t,j) \right] \right] \right.$$
$$\left. + \sum_{v=1}^{m} \left[W_v (\underline{X}_v(t) - X_v(t))^2 . \sigma(\underline{X}_v(t) - X_v(t)) \right] \right\}$$
$$+ \sum_{v=1}^{m} W_{(f)v} (\underline{X}(t_f) - X_v(t_f))^2 . \sigma(\underline{X}_v(t_f) - X_v(t_f)) \qquad (3.42)$$

where $\sigma(.)$ is the unit step function, subject to satisfaction of the difference equations for the states, (3.3) and (3.12), and where R_v is obtained from the transformation equations (3.40) and (3.41), i.e. J is a function of $U_v(t,j)$, the unconstrained variables.

3.6.4 The Hamiltonian, Gradients, and Costates

From section 3.5, J will be minimised subject to the conditions given when $\partial H(t)/\partial U_v(t,j) = 0$ for all t, j and v. The Hamiltonian, of the form of (3.25) for $t \in [1,....t_{f-1}]$ is:

$$H(t) = \sum_{j=1}^{\ell} \left\{ F\left[D(t,j) - \sum_{v=1}^{m} \left(R_v(t,j) + T_v(t) \frac{h_j}{\sum_{i=1}^{\ell} h_i} \right) \right. \right.$$
$$\left. \left. - \sum_{RR=1}^{P} G_{RR}(t,j) + L(t,j) \right] \right\}$$
$$+ \sum_{v=1}^{m} \left[W_v (\underline{X}_v(t) - X_v(t))^2 . \sigma(\underline{X}_v(t) - X_v(t)) \right]$$
$$+ \sum_{\substack{v=1 \\ v \neq q+1}}^{m} \left[\rho_v(t+1) \left(X_v(t) - \sum_{j=1}^{\ell} R_v(t,j) + I_v(t) - S_v(t) \right) \right]$$
$$+ \rho_{q+1}(t+1) \left(X_{q+1}(t) - \sum_{j=1}^{\ell} R_{q+1}(t,j) + I_{q+1}(t) - S_{q+1}(t) \right.$$
$$\left. + \sum_{j=1}^{\ell} G_q(t,j) .. k_{q/q+1} \right) \qquad (3.43)$$

(where reservoir q discharges into reservoir q+1).

And for the final time:

$$H(t_f) = \sum_{v=1}^{m} W_{(f)v}(\underline{X}_v(t_f)-X_v(t_f))^2 . \sigma(\underline{X}_v(t_f)-X_v(t_f)) \tag{3.44}$$

The costates for $v \in [1,\ldots,m]$ are:

$$\rho_v(t) = \frac{\partial H(t)}{\partial X_v(t)}$$

$$= 2W_v(X_v(t)-\underline{X}_v(t)) . \sigma(\underline{X}_v(t)-X_v(t)) + \rho_v(t+1)(1-\sigma(S_v(t))) \tag{3.45}$$

$$\rho_v(t_f) = \frac{\partial H(t_f)}{\partial X_v(t_f)}$$

$$= 2\, W_{(f)v}\,(X_v(t_f)-\underline{X}_v(t_f)) . \sigma(\underline{X}_v(t_f)-X_v(t_f)) \tag{3.46}$$

Gradients for $\left\{v \in [1,\ldots,m],\ v \neq q\right\}$ are:

$$g_v(t,j) = \partial H(t)/\partial U_v(t,j)$$

$$= \left[\lambda(t,j)\left[-1+\frac{\partial L(t,j)}{\partial G_v(t,j)}\right] - \rho_v(t+1)(1-\sigma(S_v(t)))\right] \cdot \frac{G_v(t,j)}{U_v(t,j)} \tag{3.47}$$

$U_v(t,j)$ is the unconstrained variable, which H is to be minimised with respect to

$$G_v(t,j) = R_v(t,j) + \text{constant}$$

$$\Rightarrow \frac{\partial G_v(t,j)}{\partial U_v(t,j)} = \frac{\partial R_v(t,j)}{\partial U_v(t,j)} \tag{3.48}$$

So referring to equations (3.40) and (3.41), the scaled transformations

$$\text{Region I}\quad \frac{\partial G_v(t,j)}{\partial U_v(t,j)} = 2\sqrt{MX_v(t,j)}\left[\frac{2U_v(t,j)}{\sqrt{MX_v(t,j)}} + 1\right] \tag{3.49}$$

$$\text{Region II}\quad \frac{\partial G_v(t,j)}{\partial U_v(t,j)} = 2\sqrt{MX_v(t,j)}\left[1 - \frac{2U_v(t,j)}{\sqrt{MX_v(t,j)}}\right] \tag{3.50}$$

and $$\frac{\partial L(t,j)}{\partial G_v(t,j)} = 2 \sum_{w=1}^{n} G_w(t,j) ECC_{w,v}(t,j) + ECO_v(t,j) \qquad (3.51)$$

where ECC and ECO are the appropriate sets of loss coefficients for whichever bus is slack.

Gradients for v = q, being the upper reservoir of the cascaded pair are:

$$g_q(t,j) = \partial H(t)/\partial U_q(t,j)$$

$$= \left[\lambda(t,j)\left(-1 + \frac{\partial L(t,j)}{\partial G_q(t,j)}\right) - \rho_q(t+1)\left(1-\sigma(S_q(t))\right)\right.$$

$$\left. + \rho_{q+1}(t+1)\ k_{q/q+1}\right] \cdot \frac{\partial G_q(t,j)}{\partial U_q(t,j)} \qquad (3.52)$$

where equations (3.49), (3.50) and (3.51) apply as appropriate.

This completes the model development. The equations to be solved are summarised in section 3.7

3.6.5 Justification of Spill Modelling

The amount of spill for reservoir v is:

$$S_v(t) = \begin{cases} X_v(t) - \sum_{j=1}^{\ell} R_v(t,j) + I_v(t) - \bar{X}_v \\ \text{if } \bar{X}_v < X_v - \sum_{j=1}^{\ell} R_v(t,j) + I_v(t) \\ 0 \text{ otherwise} \end{cases} \qquad (3.53)$$

No terms are included to explicitly minimise spill. This contrasts with the penalty term applied by Dillon (1974) and Dillon *et al* (1980) to impose an additional cost on spill to ensure it is minimised. At the solution spill will be minimised only insofar as unnecessarily wasted water will require extra thermal power costs in the solution. To show this, consider the Hamiltonian, eqn (3.43). At the solution:

$$\frac{\partial H(t)}{\partial U_v(t,j)} = 0 \qquad (3.54)$$

Hence either:

(a) $$\lambda(t,j)\left(-1 + \frac{\partial L(t,j)}{\partial G_v(t,j)}\right) = \rho_v(t+1)\left(1-\sigma(S_v(t))\right)$$

or (b) $$\frac{\partial G_v(t,j)}{\partial U_v(t,j)} = 0 \qquad (3.55)$$

If spilling is occurring in the t^{th} time interval then from (a) $\sigma(S_v(t)) = 1$ which implies $\lambda(t,j) = 0$ as $\frac{\partial L(t,j)}{\partial G_v(t,j)} = 1$ is not realistic. Thus in this case no thermal generation at all is required. So spilling must be an optimal strategy since there would be no demand for extra energy should the excess water be put through turbines.

Also, from equation (3.45) $\rho_v(t) = 0$ and ρ_v will be zero for all time prior to the spilling until a minimum storage constraint is violated. (Costates are calculated in reverse time order). Hence unless condition (b) applies at some time, no thermal generation will be needed over all these time intervals with $\rho_v(t) = 0$. If (b) above applies, then the maximum hydro generation possible must apply, and the excess water simply must be spilled.

3.7 SOLUTION OF THE MATHEMATICAL MODEL

The solution procedure is shown in Figure 3.5. The process begins with an initial guess of the release strategy, the present (t=1) storage in each reservoir, and the desired minimum final storage. Constrained releases are converted to unconstrained equivalents. Run-of-river station outputs are subtracted from demands, and loss coefficients are calculated for each time interval, load segment and slack bus. Then:

(i) Calculate storages (states) for each reservoir using equations (3.3) and (3.12) starting at t=1.

(ii) Equation (3.42) gives the new value of the cost function, with total transmission losses from equation (3.17).

(iii) Starting at $t=t_f$ and working backwards, costates are calculated using (3.46) and (3.45).

(iv) Gradients $\partial H/\partial U$ are found using equations (3.47) and (3.52) in conjunction with terms from (3.48) to (3.51).

(v) The conjugate gradients algorithm determines an increment to U. Provided certain convergence criteria are not met, (e.g. very small increment in U) the constrained equivalent of U is calculated by equations (3.40) and (3.41), and the process repeats from step (i) above.

3.8 CONCLUSIONS

A generalised model of a power system especially designed for the annual scheduling of hydro reservoirs has been described. Models designed for the short term optimal dispatch problem are unlikely to be suitable due to excessive detail. A solution is likely to be unobtainable when they are applied to the longer time problem. The model given avoids unnecessary detail, to help ensure the solution can be obtained reliably.

Approximations in some instances have been made that are appropriate to the New Zealand system, this particular system model being described in Chapter 5. Application to other systems would simply require different approximations to be made, about for example, the cost of thermal generation.

The solution procedure, by the Hamiltonian method, is a standard approach to dynamic optimisation. It has been described here in detail, and the relevant equations given in general form, but utilising the approximations as mentioned. State constraints are enforced by penalty functions but control constraints can be handled by either penalty functions or a transformation technique. The latter method gives exact satisfaction of constraints, but the former gives satisfaction only to within a specified tolerance level.

CHAPTER 4

THE DEVELOPMENT OF A CONJUGATE GRADIENTS ALGORITHM

4.1 INTRODUCTION

In Chapter 3 the development of a hydrothermal scheduling problem requiring a static hill climbing type method for its solution was described. The problem is summarised in section 3.7 and is solved by the conjugate gradients method of this chapter.

The use of a simple steepest descent method was not considered practicable due to the well known zigzagging difficulties it can exhibit. This behaviour can occur when a long narrow valley shape is formed by the function, and is shown for the Rosenbrock function in section 4.5.2.

Second derivative methods were rejected due to the size of the matrix to be manipulated. For eight reservoirs, weekly time intervals, and three load classes, a 1248 dimensional problem results. Operations involving each of the more than 1.55×10^6 elements in this problem's hessian will be very time consuming. Quasi-Newton methods attempt to build up an approximation to the hessian matrix, improving the estimate with the information obtained on each iteration. In theory an N-dimensional problem requires at least N iterations to obtain the exact matrix, although in practice fewer iterations may give a satisfactory solution. Likewise the 312 dimension hydro thermal problem given in Chapter 5 required between 10 and 30 iterations only, for solution by conjugate gradients (for each penalty iteration).

A serious difficulty in applying any 2nd derivative method to the problem outlined in the preceding chapter is due to the thermal cost function. This was modelled as a series of linear segments, but for the New Zealand example solution, the segments were joined by short smoothed sections. This avoids sudden changes in slope. The result is no contribution to the second derivative by fuel costs, except over the smoothing transition curves where it will have a constant value. So the hessian matrix changes suddenly but will often be near zero. A good approximation to it is obviously not going to be either obtainable, or very useful.

Conjugate gradients methods, on the other hand involve only vectors, avoiding the problems associated with handling large matrices. An apparently suitable method of this type for unconstrained problems is the I.B.M. subroutine "FMCG" which is readily available for example from Kuester and Mize (1973).

Other optimisation techniques will not be reviewed, here. A thorough review of the topic can be found in Sorenson (1976), a series of four papers covering linear programming, and constrained and unconstrained nonlinear methods.

After giving a general conjugate gradients formula, the simplified form used by Fletcher and Reeves is given in section 4.2.3. Their algorithm has difficulty handling non-linearities, so a generalised version is described in detail in section 4.4. The results of various tests are presented in section 4.5. Results are given for the solution of a two-dimensional Rosenbrock function which has smooth contours, but is non-convex. Our method (called "CGRADS" from Conjugate GRADientS), was developed to handle the difficult, non-linear features of the hydro-thermal problem. Performance for this severe test is given for the three algorithms compared - steepest descent, FMCG and CGRADS.

4.2 A FORMULA FOR CONJUGATE GRADIENTS

4.2.1 Definition and Justification

The justification for conjugate gradient algorithms comes from consideration of the quadratic function:

$$F(\underset{\sim}{x}) = \frac{1}{2} \underset{\sim}{x}^T A \underset{\sim}{x} + \underset{\sim}{b}^T \underset{\sim}{x} + C \qquad (4.1)$$

where $\underset{\sim}{x}$ is an n-dimensional vector.
A set of directions $\underset{\sim}{d}_i$ (i=1,2...) are said to be mutually conjugate with respect to A if

$$\underset{\sim}{d}_i^T A \underset{\sim}{d}_j = 0 \quad \text{for } i \neq j \qquad (4.2)$$

If the directions are non-null, and are used for successive searches, with F(x) being minimised along each, then (from Beale, 1972):

(a) After any number, r, of steps, we have the minimum value of F in the hyperplane spanned by the search directions used.

(b) After at most n steps, the unconstrained optimum will be found.

Fletcher and Reeves (1964) prove how a set of conjugate search directions minimise a function, but earlier, Beckman (1962) had showed how they can be used to solve a set of simultaneous linear algebraic equations.

4.2.2 A General Formula

Beale (1972) describes how to generate a set of conjugate directions, beginning with an arbitrary, downhill direction. The resulting process is:

$$\begin{aligned} \underset{\sim}{d}_1 &= \text{arbitrary downhill direction} \\ \underset{\sim}{d}_2 &= -\underset{\sim}{g}_2 + \beta_2 \underset{\sim}{d}_1 \\ \underset{\sim}{d}_k &= \underset{\sim}{g}_k + \beta_k \underset{\sim}{d}_{k-1} + \gamma_k \underset{\sim}{d}_1 \qquad (k>2) \end{aligned} \tag{4.3}$$

where $$\beta_k = \frac{\underset{\sim}{g}_k^T(\underset{\sim}{g}_k - \underset{\sim}{g}_{k-1})}{\underset{\sim}{d}_{k-1}^T(\underset{\sim}{g}_k - \underset{\sim}{g}_{k-1})} \tag{4.4}$$

$$\gamma_k = \frac{\underset{\sim}{g}_k^T(\underset{\sim}{g}_2 - \underset{\sim}{g}_1)}{\underset{\sim}{d}_1^T(\underset{\sim}{g}_2 - \underset{\sim}{g}_1)} \tag{4.5}$$

and $\underset{\sim}{g}_k$ is the gradient at the beginning of the k^{th} search.

After n-iterations there is no mathematical justification for continuing to use these equations, for an n-dimensional quadratic function. So the process should be restarted, the n^{th} search direction ($\underset{\sim}{d}_n$) becoming the first member of the new set, i.e. $\underset{\sim}{d}_1$.

For a general function, where eqn (4.1) represents only the first three terms of a Taylor series approximation, A will not be a constant. Restarting more frequently is advantageous. Powell (1977) suggests restarting whenever the following inequality is satisfied:

$$\left|\underset{\sim}{g}_{k-1}^T \underset{\sim}{g}_k\right| > \text{RSC}\left|\underset{\sim}{g}_k\right|^2 \tag{4.6}$$

where RSC = 0.2 or 0.5 appears satisfactory. This provides a means for determining when the directions being calculated are no longer near enough to conjugate. For example, the second derivative matrix A might have changed too much.

A restart effectively discards accumulated information and starts afresh. Restarting with a steepest descent step could lead to trouble as the function value reduction is likely to be less than with a conjugate

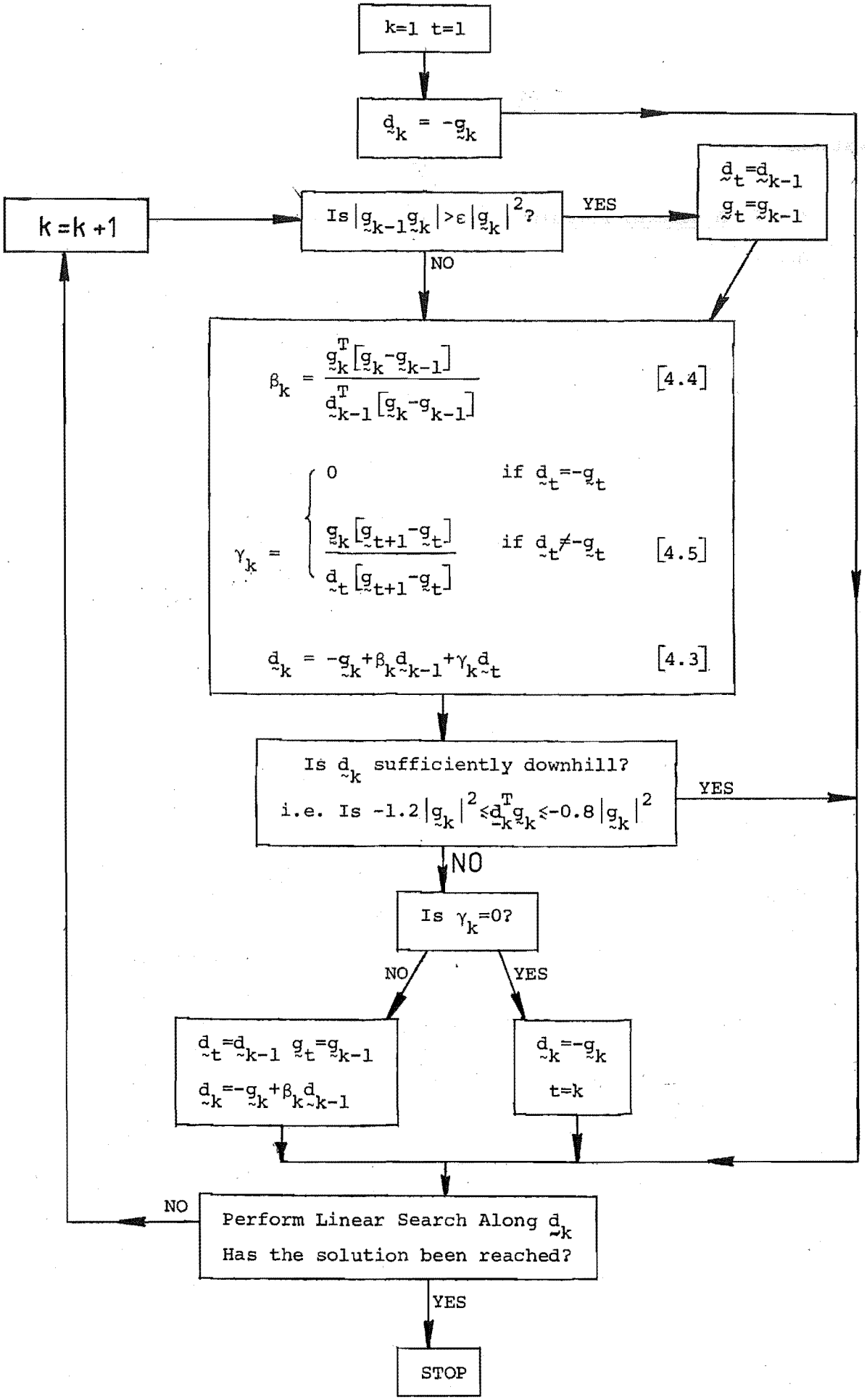


FIGURE 4.1 Calculation Flow Chart for Conjugate Gradients

directions step. Using the Beale formulas (eqns (4.3) to (4.5)) allows the previous search direction to become the first of the new series. So the information in this step is utilised in calculating a conjugate direction immediately, rather than having to take a steepest descent step first.

One further check is needed on search directions - that they are sufficiently downhill. From Powell (1977) again, a restart is made if the inequalities are not satisfied:

$$-1.2\ |g_k|^2 \leq d_k^T g_k \leq -0.8|g_k|^2 \qquad (4.7)$$

The static optimisation algorithm, for which more details are given in section 4.4, used to solve the hydro-thermal problem, is summarised in Figure 4.1. It shows how restarts and search direction checks fit into the process.

4.2.3 A Simplified Conjugate Gradients Formula

The equations (4.3) and (4.4) can be simplified by making some approximations to give the formulation for conjugate gradients of Fletcher and Reeves.

(i) Exact Linear Searches: If each linear search were to locate the true minimum along the direction d_k then:

$$d_k \cdot g_{k+1} = 0 \qquad (4.8)$$

i.e. the gradient at the minimum is orthogonal to the search direction. The denominator for β_k becomes:

$$\begin{aligned} & d_{k-1} \cdot (-g_{k-1}) \\ = \ & (-g_{k-1} + \beta_k d_{k-2}) \cdot (-g_{k-1}) \\ = \ & g_{k-1} \cdot g_{k-1} \end{aligned}$$

Hence:
$$\beta_k = \frac{g_k^T [g_k - g_{k-1}]}{|g_{k-1}|^2} \qquad (4.9)$$

which is used in the Polak Ribiere (1969) conjugate gradients algorithm.

(ii) Quadratic function: If the quadratic function of eqn (4.1) is to be minimised, then the gradients $\underset{\sim}{g}_1$, $\underset{\sim}{g}_2$,...$\underset{\sim}{g}_k$ are mutually orthogonal. (See Beale (1972) for a proof). Then β_k in eqn (4.9) can be further reduced to

$$\beta_k = |\underset{\sim}{g}_k|^2 / |\underset{\sim}{g}_{k-1}|^2 \qquad (4.10)$$

(iii) Initial Search: If the initial search is in the direction of steepest descent, then γ_k of eqn (4.5) is zero. Hence Fletcher and Reeves formula

$$\underset{\sim}{d}_k = -\underset{\sim}{g}_k + \beta_k \underset{\sim}{d}_{k-1} \qquad (4.11)$$

where β_k is given in eqn (4.10).

4.3 FLETCHER AND REEVES CONJUGATE GRADIENTS

I.B.M. subroutine FMCG implements the algorithm described in Fletcher and Reeves (1964) paper, and so uses eqns (4.10) and (4.11). The flow chart, Fig. 4.2 shows the procedure. A steepest descent step is taken on the first iteration, and after every N+1 iterations, for an N dimensional problem. During the linear search checks are made for a positive directional derivative and any increase in function value. If either is detected, cubic interpolation is used. Interpolation is repeated on successively smaller intervals until a value less than either of the end points is found.

A single parameter, EPS, is then used to check for convergence. This is clearly unsatisfactory as the value of GNRM = $\sum_{i=1}^{N} g(i)^2$ (length of gradient vector squared) bears no relation to the sum of changes in state which is small enough to indicate convergence. Likewise, the tolerable increase in function value over an iteration will be different again. Test runs made used three separate parameters (EPSF, EPSX, and EPSG) to check these quantities, instead.

Dimension of Problem = N.

KOUNT = Iteration counter.

F = Function value.

X = N-vector of states.

G = N-vector of gradients.

SD = 2N-vector. 1st N elements are the Search Direction.

DY = Direction Derivative.

$= \underset{\sim}{G}.\underset{\sim}{SD}.$

$= \sum_{i=1}^{N} G(i).SD(i).$

$= |G|.|SD| \cos\theta.$

SNRM = Step size.

GNRM $= |G|^2 = \sum_{i=1}^{N} G(i)^2.$

OLDG = GNRM for previous iteration.

EPSF = Absolute amount F may increase by over an iteration.

EPSX = Convergence tolerance on state variables.

EPSG = Convergence tolerance on gradients but where EPS = EPSG = EPSX for standard IBM version.

IER = error parameter

1 Max. no. of iterations (LIMIT) exceeded.

2 No minimum exists.

0 Converged.

-1 error in gradient calculations.

FIGURE 4.2(a) Fletcher and Reeves Conjugate Gradients Algorithm - As Implemented by I.B.M. Program FMCG.

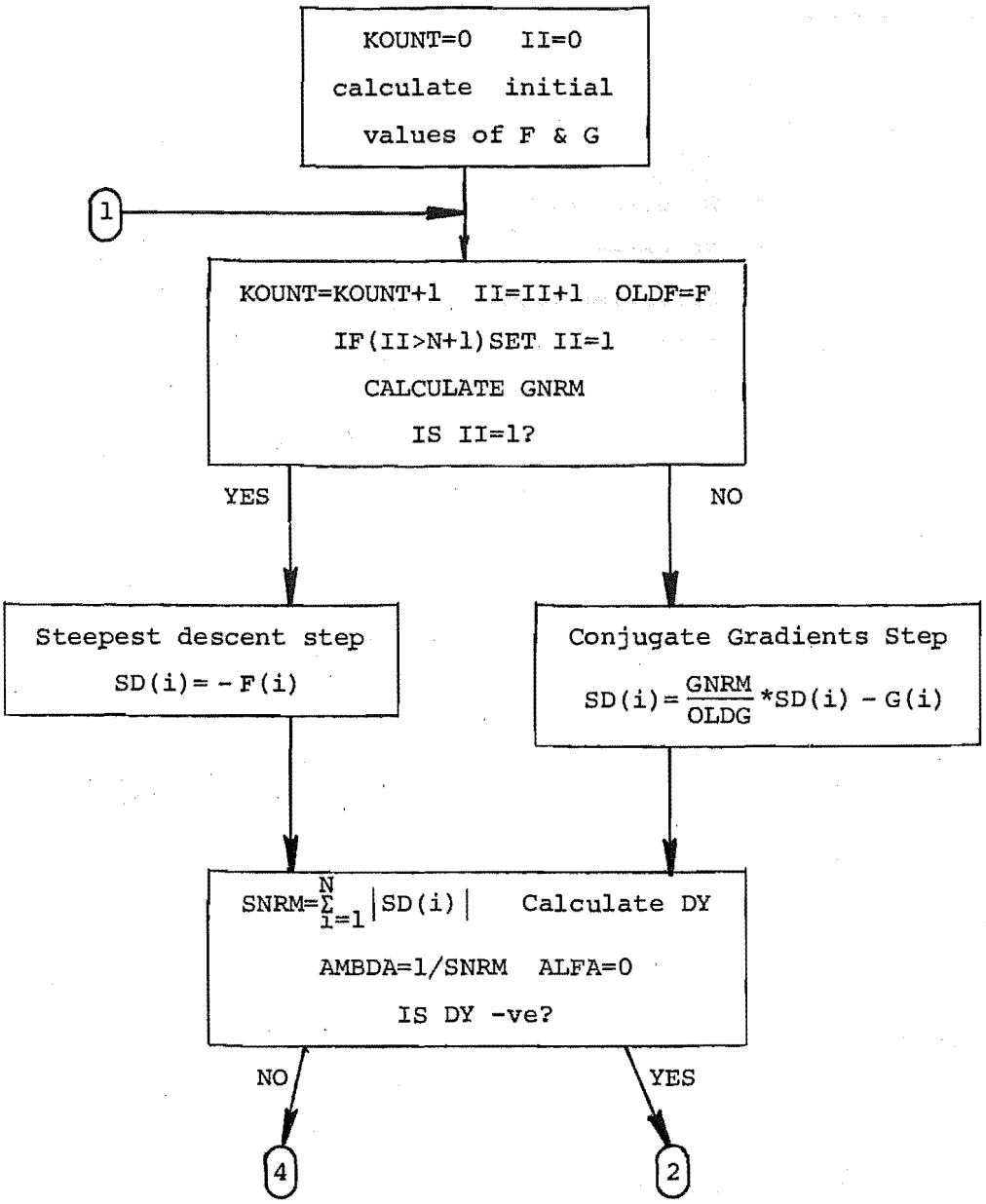


FIGURE 4.2(b) Fletcher and Reeves Conjugate Gradients

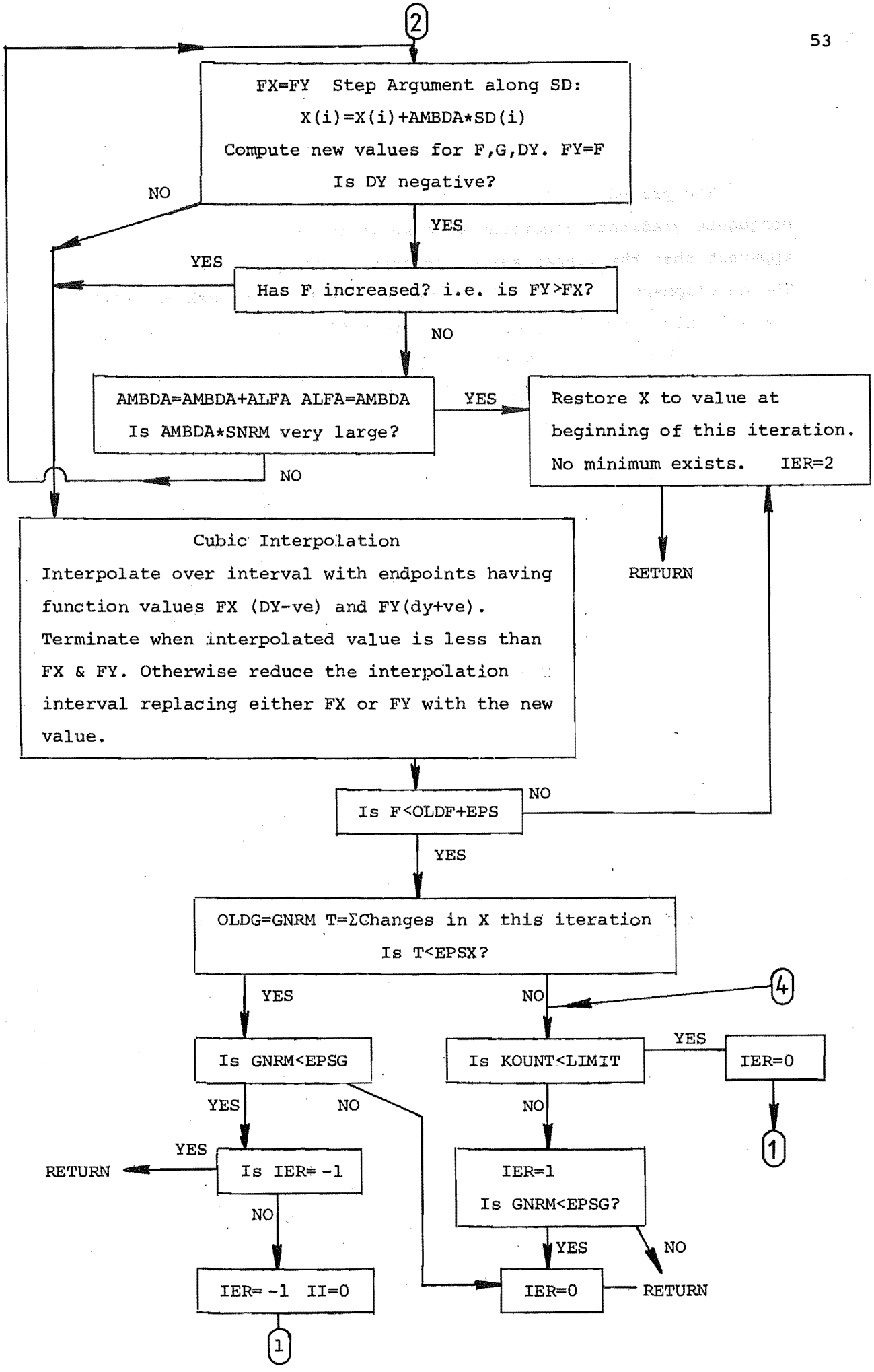


FIGURE 4.2(c) Fletcher and Reeves Conjugate Gradients

4.4 A RELIABLE CONJUGATE GRADIENTS ALGORITHM

4.4.1 Introduction

The procedures for determining search directions for a generalised conjugate gradients algorithm were shown in Figure 4.1. It soon became apparent that the linear search process of FMCG was unsatisfactory. The development of a suitable linear search strategy making CGRADS a reliable algorithm is described. First the reasons for the stepsize adjustment strategy are given, then an accuracy requirement, and then the difficulties leading to the use of gradients only to control the search. Finally the various tests needed under different circumstances to decide when to terminate the search are described. Figure 4.4 shows details of CGRADS.

4.4.2 Linear Search Stepsize

The initial step size along a search direction is calculated by both FMCG and CGRADS as

$$\text{Initial stepsize} = \frac{1}{\sum_{i=1}^{N} SD(i)} \qquad (4.12)$$

This is somewhat arbitrary. Quasi Newton methods have a step size which tends to 1.0 as the solution is approached, (see Shanno, 1978). However no step size information is available for conjugate gradient methods. So the initial step could even result in an immediate increase in function value. FMCG has no provision for trying a smaller step, and would use cubic interpolation in this situation. CGRADS can adjust its step size by doubling it or by shortening to examine the midpoint of the interval within which the minimum is known to exist. This ability to reduce step-sizes below their initial values can be especially important near the solution.

4.4.3 Accuracy of Linear Searches

FMCG has no accuracy requirements imposed on its linear search - the minimum need only be bracketed and one successful cubic interpolation carried out. The approximation to simplify eqn (4.4) to get eqn (4.9) required that an exact minimum be found along the search direction. This would give a zero directional derivative eqn (4.8).

To help ensure that search directions are approximately conjugate, a specific reduction in directional derivative over each linear search is required:

$$XLS \cdot \underset{\sim}{d}_i \cdot \underset{\sim}{g}_i = \underset{\sim}{d}_i \cdot \underset{\sim}{g}_{i+1} \qquad (4.13)$$

where: XLS is typically 0.05 to 0.2, specifying the required reduction in directional derivative

$DY = \underset{\sim}{d}_i \cdot \underset{\sim}{g}$

$\underset{\sim}{g}_i$ = gradient at beginning of linear search

$\underset{\sim}{g}_{i+1}$ = gradient at end of linear search.

This strategy ensures that departure from conjugacy is slowed.

4.4.4 Controlling the Linear Search

A minimum along a linear search should be bracketed by two points with directional derivatives, DY, of opposite sign. Should the function not be unimodal along the search direction, the function value could increase while DY remains negative. Figure 4.3 gives an example of the contours that could be produced by the switching in of another thermal station, in the hydro-thermal problem. As a new set of D.C. loadflow loss coefficients is then needed, a jump in transmission losses is possible, due to the approximations in the loss model.

So function values could suggest a minimum has been passed, but DY indicates otherwise. CGRADS examines only DY values, to ensure such "glitches" in contours do not cause premature termination of the linear search. Even if the function value at B in Figure 4.3 were larger than that at A (by the discontinuity), termination at B is thought to be preferable - a nasty discontinuity is not a good place to begin the next iteration from. Convergence is likely to be aided by avoiding such points.

Earlier versions of CGRADS did examine function values, but a dramatic improvement in convergence reliability resulted from using DY values only, to control the linear search.

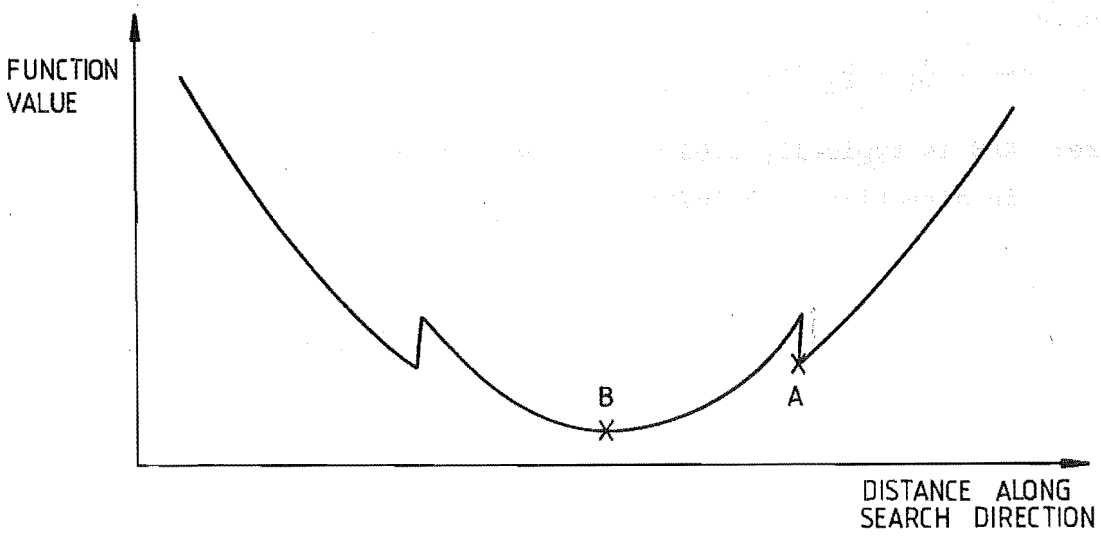


FIGURE 4.3 "Glitch" in contours of the hydro-thermal scheduling problem

4.4.5 Terminating the Search

Three conditions are tested to assess a linear search's effectiveness, see Figure 4.4(a):

(i) DY is reduced sufficiently (eqn 4.13).

(ii) The function value is less than at the beginning of that iteration, i.e. F<FI.

(iii) The function value is less than at the beginning of the last successful iteration, i.e. F<FOLD.

Some limit must be imposed on the smallest possible step size adjustment. Round-off errors in the computer impose a lower limit, but the shape of the function contours may make condition (i) difficult to meet. e.g. A large second derivative gives a rapidly changing gradient, and so the required value to gradient to satisfy (i) is difficult to find. Imposing suitable lower limits on step size adjustments can avoid wasted computation around difficult contours. If this limit is reached without satisfying condition (i), (ii) and (iii) are checked. If both are met, the search is successful. If only (ii) is met, then it is assumed that the solution has been reached, i.e. no further reduction in F is possible.

Dimension of Problem = N.

KOUNT = Iteration Counter.

F = Function Value.

FI = Function Value at beginning of this iteration.

FOLD = Function value at beginning of previous iteration.

X = N vector of states.

G = N vector of gradients.

SD = 2N vector. 1st N elements are the search direction. 2nd N elements are the values of states at the beginning of the iteration.

GL = N vector of gradients at beginning of linear search process.

GDRS = GL of previous iteration, stored when restarting.

SDRS = SD of previous iteration stored when restarting.

DY = Directional derivative = $\underset{\sim}{G}.\underset{\sim}{SD}$

SZ = Step size.

GNRM = $\sum_{i=1}^{N} G(i)^2$

LSZ = Largest step size so far with negative DY.

HSZ = Smallest step size so far with positive DY.

Four types of search possible

II=1	RS=1	steepest descent.
II=1	RS=0	steepest descent started series conjugate direction.
II=0	RS=1	Beale restart direction.
II=0	RS=0	Beale restarted series conjugate direction.

FIGURE 4.4(a) Flow Chart for Generalised Conjugate Gradients Algorithm "CGRADS".

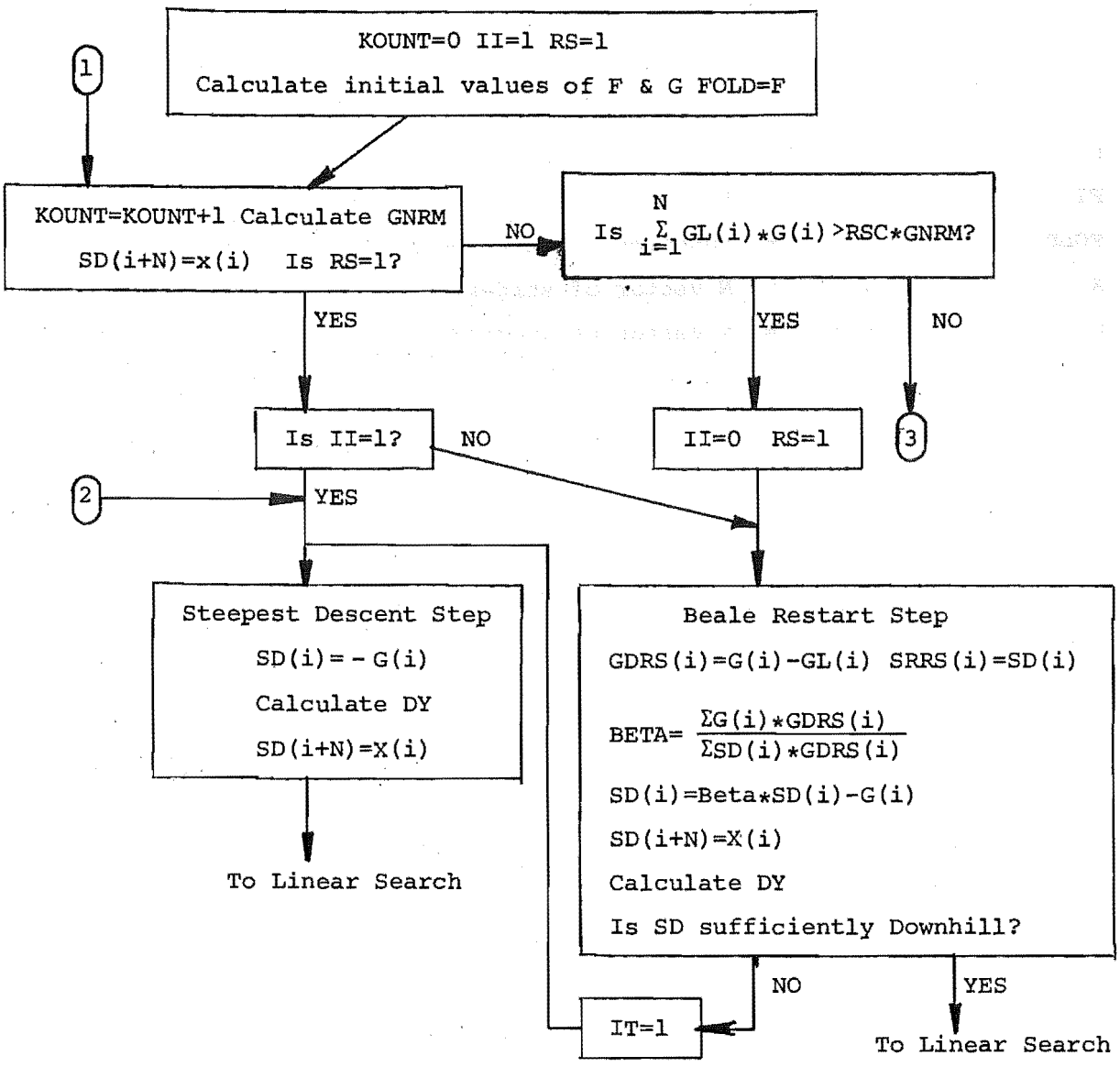


FIGURE 4.4(b) Generalised Conjugate Gradients "CGRADS"

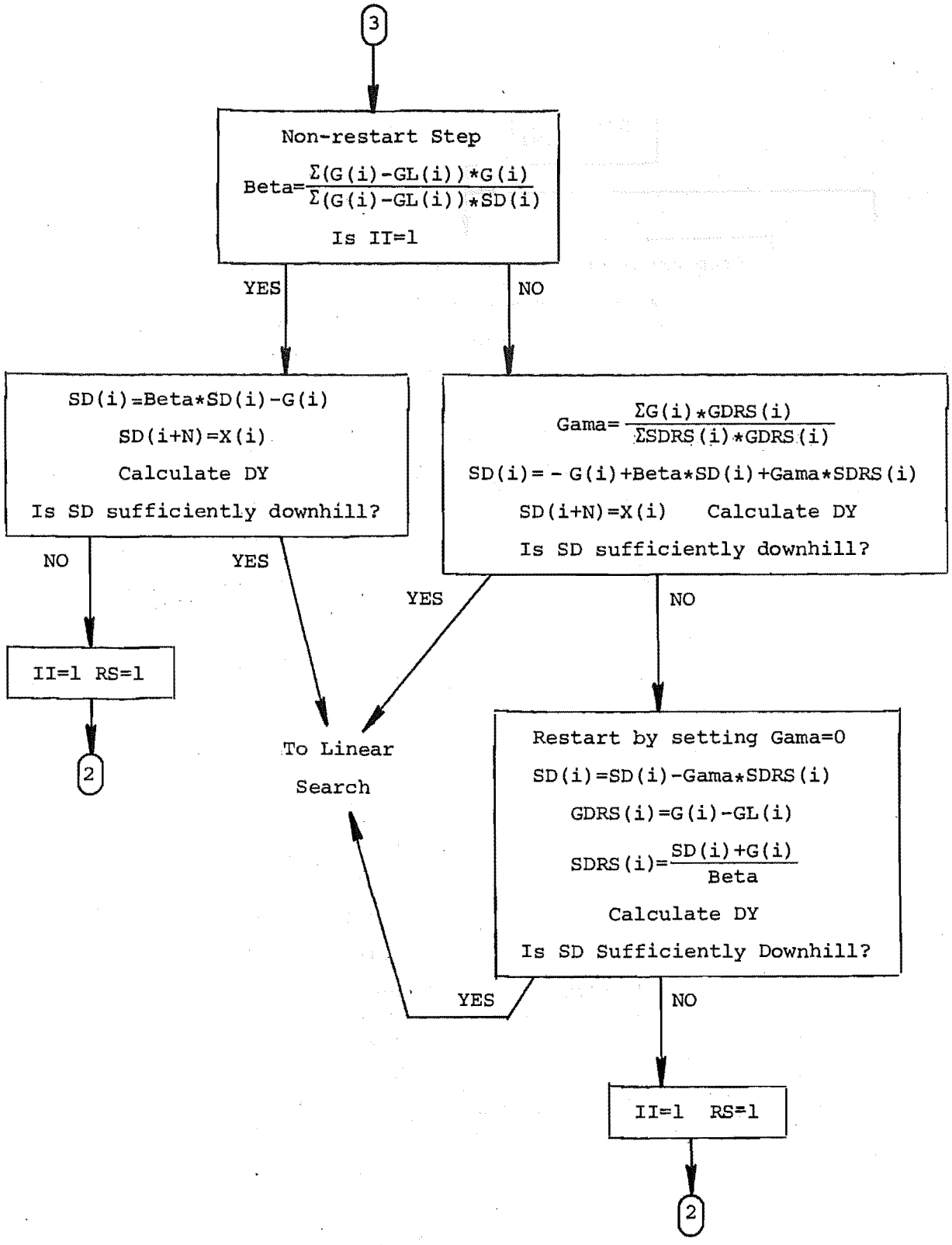


FIGURE 4.4(c) Generalised Conjugate Gradients "CGRADS"

Linear Search

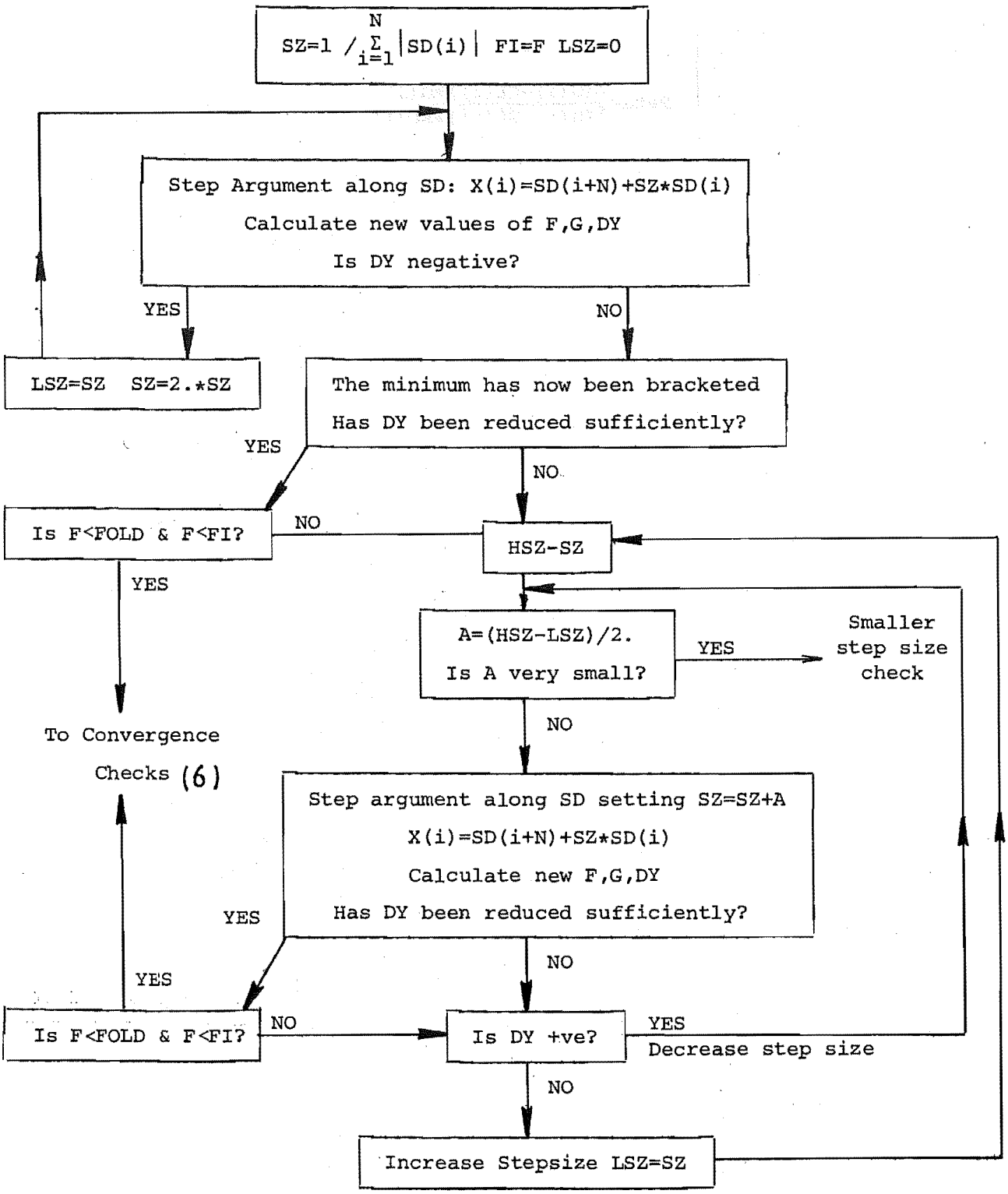


FIGURE 4.4(d) Generalised Conjugate Gradients "CGRADS"

Smaller Step Size Check

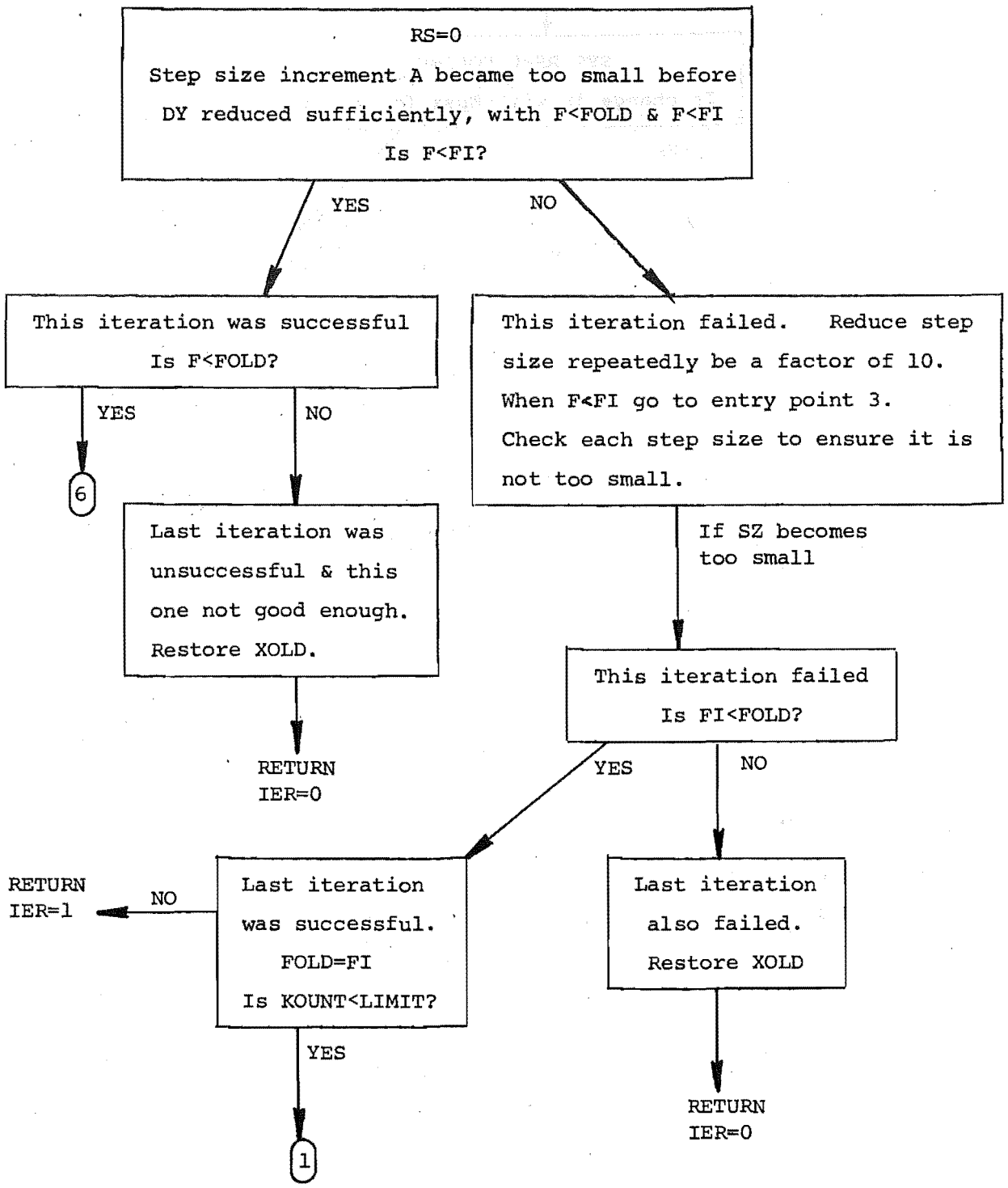


FIGURE 4.4(e) Generalised Conjugate Gradients "CGRADS"

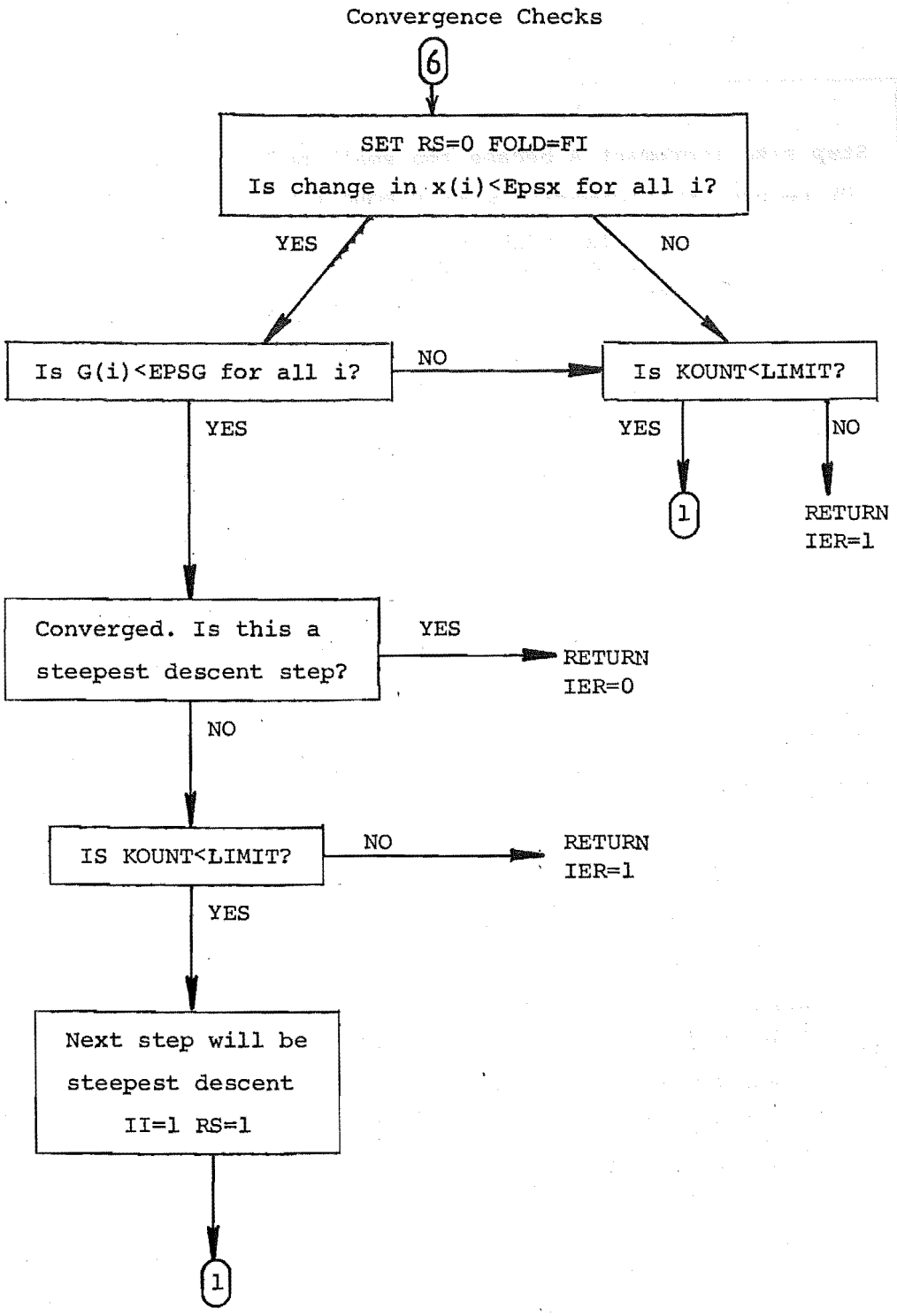


FIGURE 4.4(f) Generalised Conjugate Gradients "CGRADS"

If (ii) is not met - no reduction in function value on this iteration - then a check is made by reducing the step size by 10 repeatedly. This is to ensure that the initial step did not completely overshoot the region of the minimum. An example of this phenomenon was observed when solving the two dimensional Rosenbrock function (see Figure 4.5):

$$f(x_1,x_2) = 100.\ (X_2-X_1^2)^2 + (1-X_1)^2 \qquad (4.14)$$

As it is a non-convex function, a large enough initial step could, for some search directions, shoot the search from the bottom of one end of the valley, at A, right across a region of higher function values, to B, at the other end of the valley. It might not be possible to reduce the function value at this end of the valley below its initial value. The reduced stepsize check enables recovery from this situation.

No reduction in F might result from this check, in which case the previous iteration is looked at. If it was also a failure (i.e. FI>FOLD) then it is assumed that the solution has been reached. Otherwise, the process carried on with another iteration.

The objective of all these checks etc. is to avoid hanging up on awkward contours - it is better to accept a few inconsistencies and carry on if possible.

Convergence checks are made to test whether another iteration is needed (see Figure 4.4(f)). The absolute value of the change in each state variable is tested, not the sum of all changes. A test of the sum might not recognise when a few variables are still changing significantly. Similarly, the values of each gradient term are checked. If both tests are passed then a steepest descent step must be performed, and convergence checked again, before terminating. This is to ensure that the conjugate gradients formulas have not produced strange search directions permitting only small changes in state. As it is difficult to decide what is a reasonable value for the gradient at the solution, only a coarse check can be made on gradients to help show when something has gone seriously wrong.

4.5 TESTING THE ALGORITHMS

4.5.1 The Rosenbrock Function

The two dimensional Rosenbrock function (eqn. (4.14)) is shown in Figure 4.5. At the solution point (1., 1.), its value is 0. Being long, narrow and curved, it gives the type of contours that a steepest descent algorithm has difficulty with.

The steepest descent method's failure is shown in Figure 4.6. After about 500 function evaluations, the function is almost down to 10^{-2}, but a further 4,500 evaluations produce little improvement - a severe case of zigzagging.

FMCG is much more successful. In Figure 4.7 it has reduced the function to 10^{-9} after only 250 function evaluations and to 10^{-11} after a total of 470. Clearly, the principle of conjugate directions really does work, and gives a vast improvement on steepest descent.

CGRADS reduces the function to 10^{-9} after 175 function evaluations, but goes on to reduce it to 10^{-23} after 400 evaluations, as in Figure 4.9. Both conjugate directions algorithms are effective, but the accuracy of FMCG is limited to a minimum function value of 10^{-11}. No further reduction could be obtained by using smaller convergence tolerances. This is due to the arbitrary initial stepsize used, with no mechanism for its reduction.

If FMCG is run beginning each linear search with a step size 10^{-3} times that normally used, the results in Figure 4.8 are obtained. A very slow reduction in function value to 10^{-2} occurs over the first 225 function evaluations, then a sudden drop to 10^{-28} over the next 35 evaluations. Good performance on this problem by FMCG is to be expected - it has no non-linearities or discontinuities, and the cubic interpolation is most effective on the quadratic Rosenbrock function.

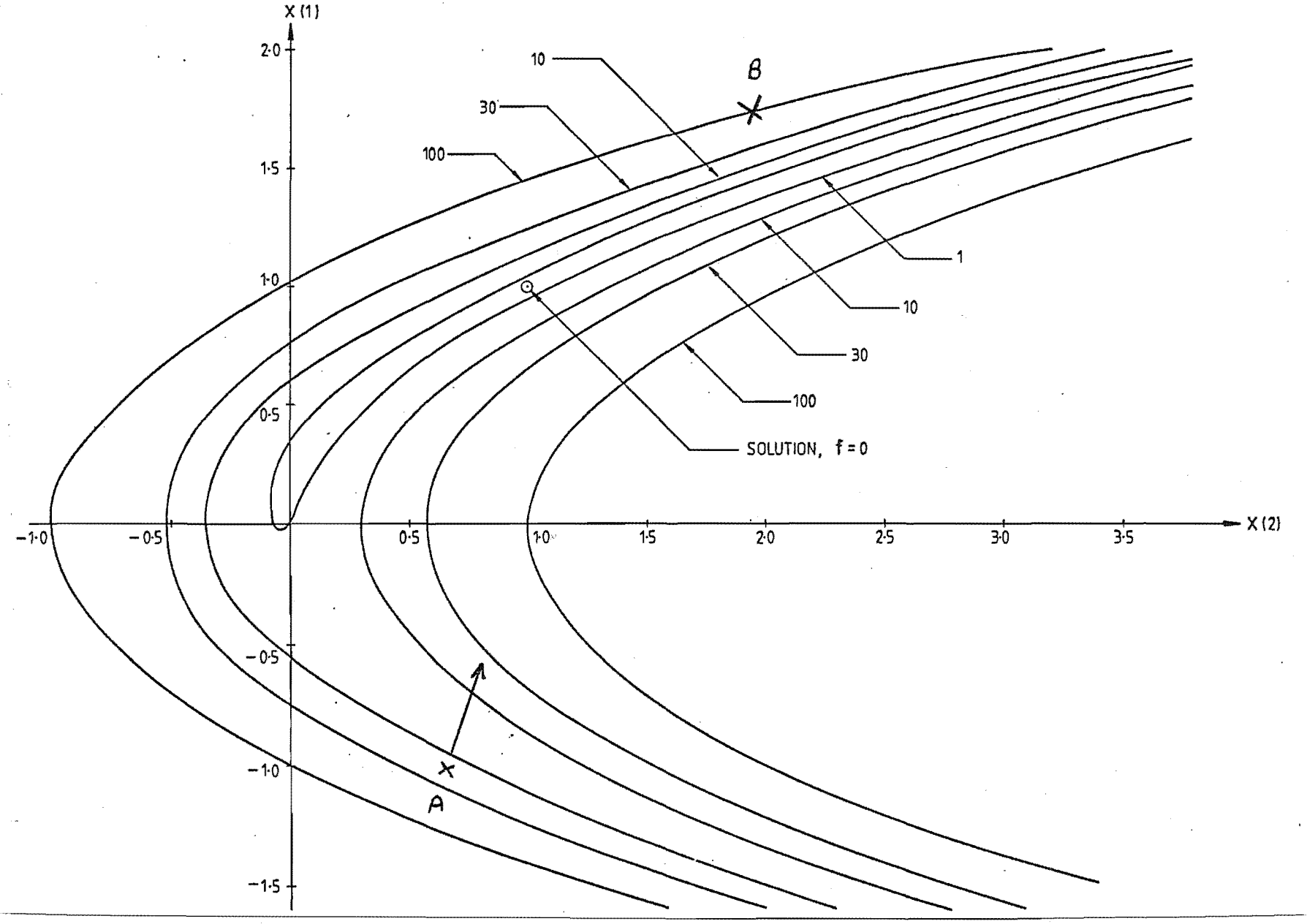


FIGURE 4.5 2-Dimensional Rosenbrock Function, $f(x_1, x_2)=100.(X_2-X_1^2)^2 + (1-X_1)^2$

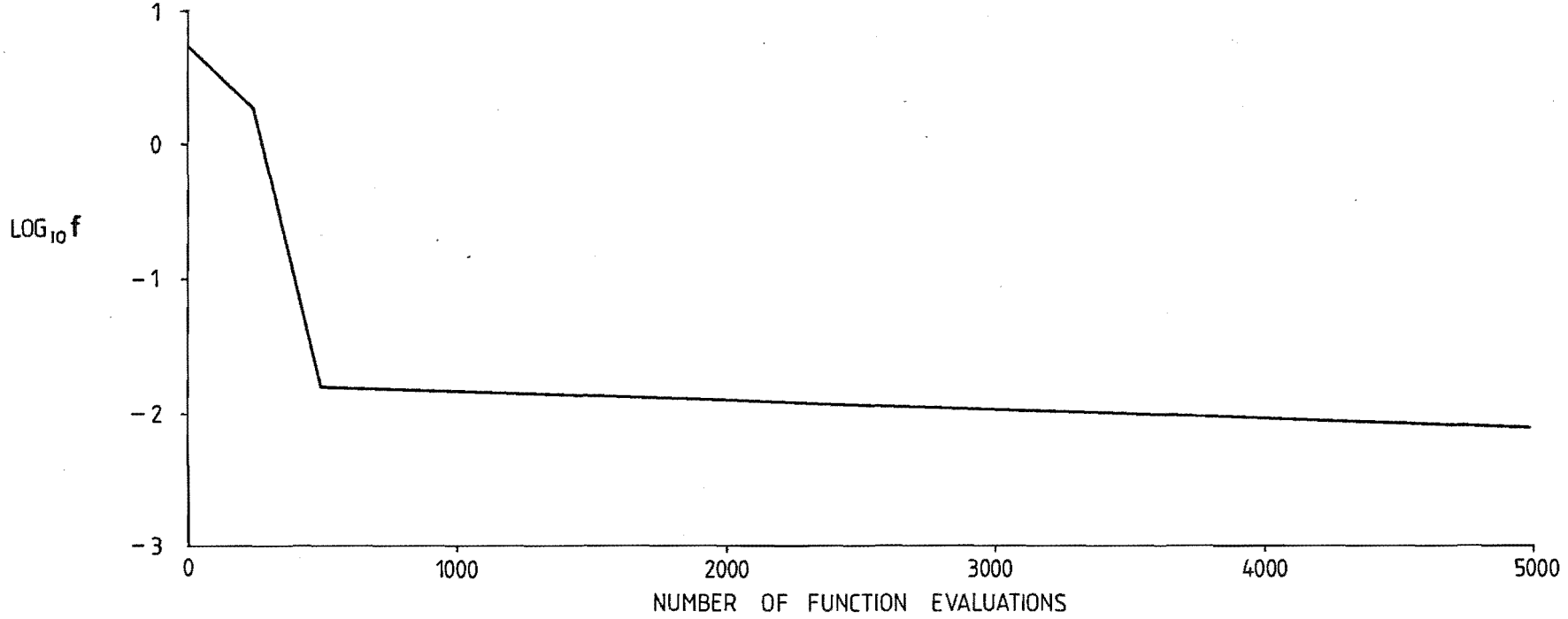


FIGURE 4.6 Reduction in function value, Steepest Descent for 2-Dimensional Rosenbrock function.

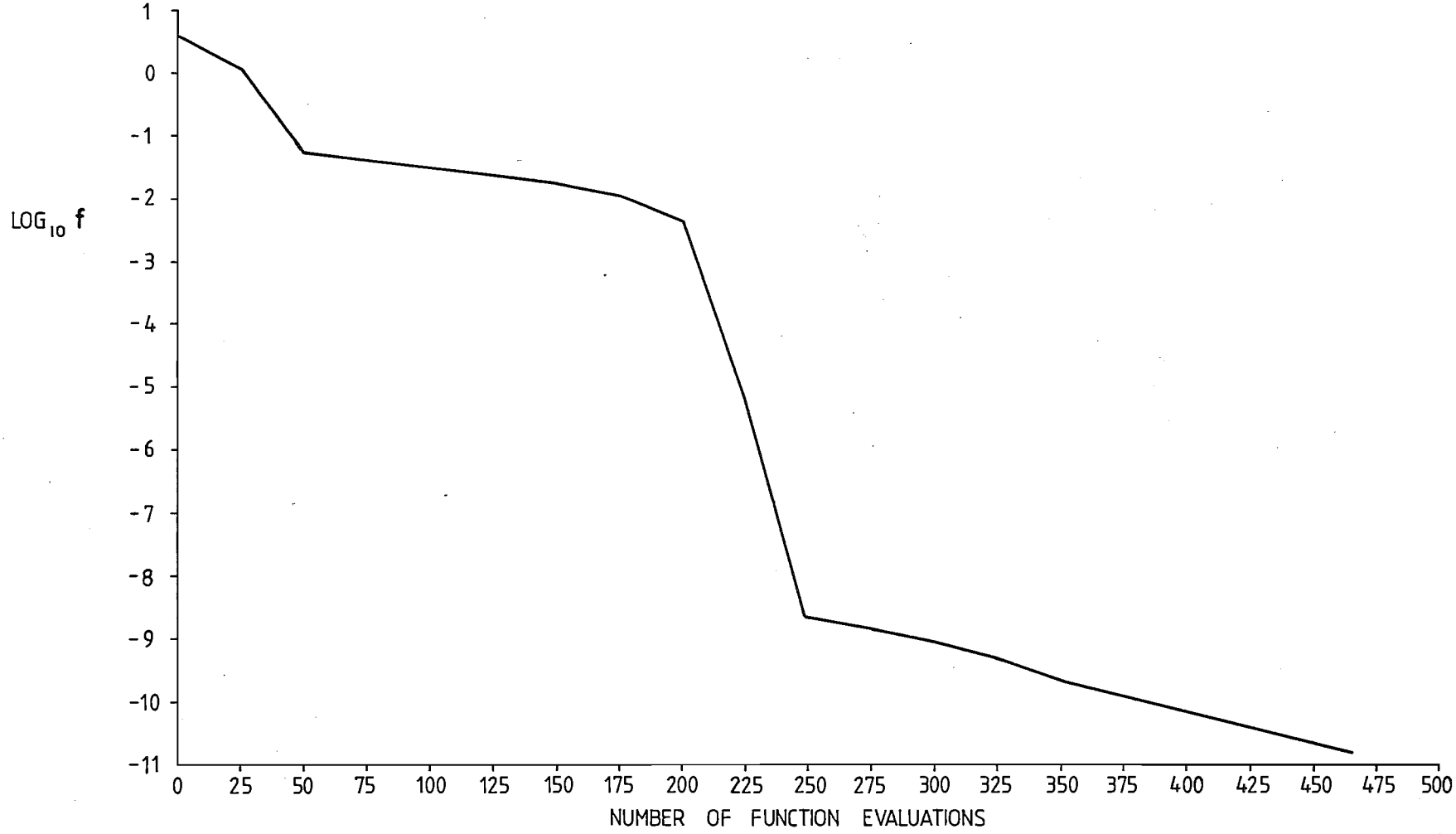


FIGURE 4.7 Reduction in function value, FMCG for 2-Dimensional Rosenbrock function.
(xl. scaling of initial stepsize)

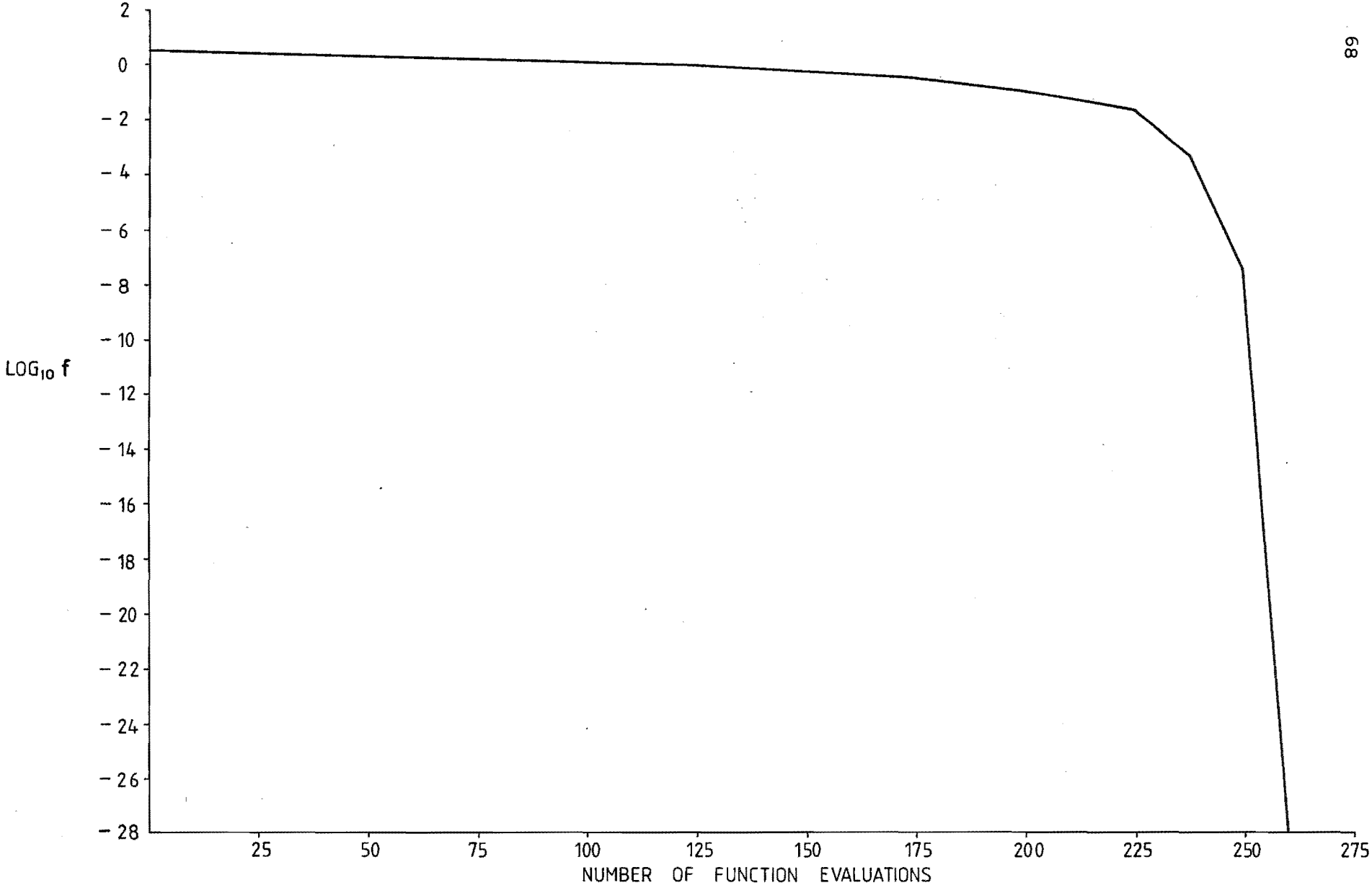


FIGURE 4.8 Reduction in function value, FMCG for 2-Dimensional Rosenbrock function (x.001. scaling of initial stepsize).

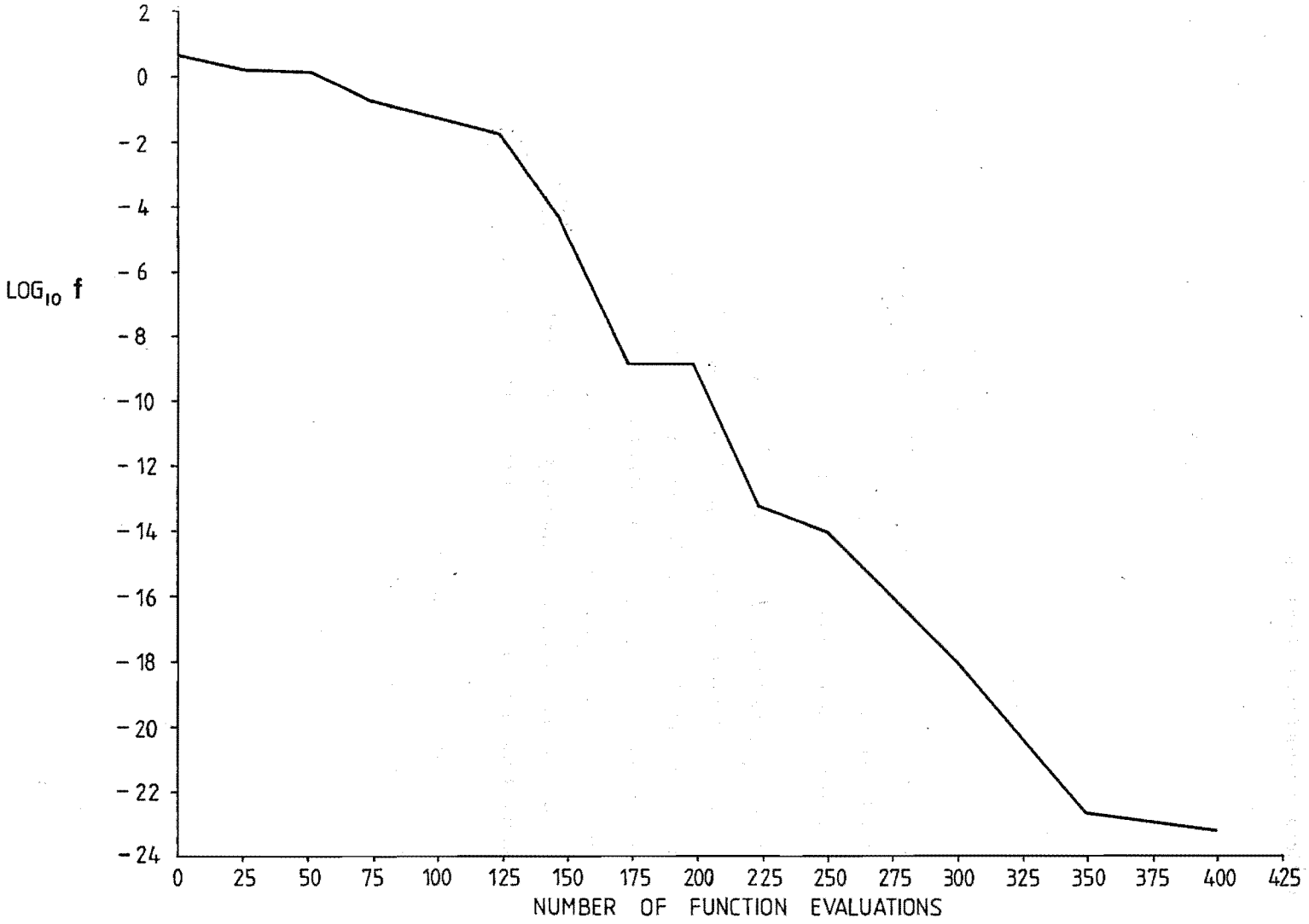


FIGURE 4.9 Reduction in function value, CGRADS for 2-Dimensional Rosenbrock function.

4.5.2 Peformance for Hydro-Thermal Scheduling Problems

CGRADS has been used to solve the eight reservoir, 312 dimension hydro thermal problem with three different constraint enforcement strategies:

(i) penalty functions enforcing both state and control constraints;

(ii) unscaled control transformations, state penalties;

(iii) scaled control transformations, state penalties, and with a variety of penalty function powers.

One reservoir stochastic linear feedback models solved include constant gain, variable gain and gaussian plus impulse versions.

The success of CGRADS on these problems gives confidence in its reliability. This is a more important attribute than economy of processing, under the computer facilities used, as computer time is not charged to the user.

FMCG was the least reliable of the three hill-climbing techniques tested. It succeeded only with the scaled control transformations version, requiring 265 minutes processing and 33,000 function evaluations.

The steepest descent method was more successful. Results are summarised in Table 4.1 with more details in Figures 4.10 and 4.11. The state and control penalty functions version required 24 minutes (11 minutes for CGRADS), but the unscaled problem 129.5 minutes or 9 times as long as CGRADS. The factor of only two saving in using CGRADS on the scaled problem, as against 9 on the unscaled, is to be expected. The latter has a greater likelihood of striking long narrow contours which can be handled easily by CGRADS

Table 4.2 gives some results for CGRADS, showing the number of iterations of each of the four types. Steepest descent steps give the greatest function value reductions, because the first iteration, having the greatest potential for function value reduction, is always in the direction of steepest descent. In two cases conjugate gradient (0,0) type steps give the smallest function reductions.

The order in which the various types of steps are taken can be seen from Table 4.3, for the first of the four cases summarised in Table 4.2. The nature of contours is likely to change so much over the first two steps that no information of further use is gained - hence a conjugate directions step is not taken until the third iteration. A series of seven conjugate directions steps follows before a restart on the ninth step.

Towards the end of the penalty iteration, more frequent steepest descent restarts are made. This is because the last step must be a steepest descent type. Whenever convergence tolerances are met, a steepest descent step is taken and convergence rechecked.

Referring back to Table 4.1, satisfactory solutions can be obtained for the scaled problem with XLS (linear search accuracy) ranging from 1.0 to 0.05, with RSC (restart parameter) taking a similar range. The unscaled form is more sensitive, XLS = 0.2 being more efficient than the other values shown. RSC values outside the range 0.2 to 0.5 also gave less accurate solutions, for some XLS values, even when convergence tolerances were reduced.

For difficult problems, where accurate solutions were important, the combination XLS = 0.05, RSC = 0.5 was found to be most successful. It was used for the various penalty function power tests and the Gaussian plus impulse models.

TABLE 4.1 Results for Deterministic 8 Reservoir Hydro Thermal Problem, for Various XLS (Linear Search Accuracy) and RSC (Restart Parameter).

Top figure in each square = computer process time, minutes.
Bottom figure = cost of thermal fuel, dollars, to a base of $\$0.235 \times 10^8$.

Unscaled Control Transformations:

XLS \ RSC	CGRADS 1	CGRADS 0.5	CGRADS 0.2	CGRADS 0.05	Steepest Descent
1.	43.7 1.011	40.9 1.007	40.4 1.016	40.2 1.014	143.5 1.004
0.5	73.9 1.003	80.8 1.002	77.8 1.007	48.4 1.008	129.5 1.004
0.2	27.1 1.008	15.4 1.002	21.2 1.006	16.8 1.001	208.0 1.008
0.05	34.0 1.020	36.8 1.000	64.3 0.998	29.2 1.043	- -

Scaled Control Transformations

XLS \ RSC	CGRADS 1	CGRADS 0.5	CGRADS 0.2	CGRADS 0.05	Steepest Descent
1.	12.7 0.998	12.6 0.998	10.0 1.002	11.7 0.997	25.2 0.997
0.5	10.6 0.997	12.0 1.000	12.5 1.000	11.6 0.998	21.8 0.998
0.2	12.7 0.998	12.4 1.000	15.7 1.000	12.5 1.001	42.7 0.998
0.05	16.3 1.000	18.3 0.997	18.5 0.999	19.1 0.994	- -

TABLE 4.2

Problem: Scaled Control Constraints, Deterministic Hydro-Thermal.

Algorithm: CGRADS (XLS = 0.2, RSC = 0.5)

Also see Figure 4.12

Iteration Type		Number of Iterations	Av. Function Evals Per Iteration	Av. %age Function Change Per Iteration	Av. %age Function Change Per Function Evaluation
2nd Penalty Iteration:					
1	1	6	6	4.52	0.75
1	0	12	12	3.21	0.27
0	1	13	8	0.17	0.021
0	0	20	10	0.010	0.001
3rd Penalty Iteration:					
1	1	2	10	12.0	1.20
1	0	12	9	2.56	0.28
0	1	5	12	0.41	0.034
0	0	3	19	2.08	0.11

Problem: Unscaled Control Constraints, Deterministic Hydro-Thermal.

Algorithm: CGRADS (XLS = 0.2, RSC = 0.5)

Also see Figure 4.13

2nd Penalty Iteration:					
1	1	6	4	1.07	0.27
1	0	52	4	0.686	0.17
0	1	12	3	0.524	0.17
0	0	16	4	0.972	0.24
3rd Penalty Iteration:					
1	1	3	7	3.91	0.56
1	0	9	5	3.84	0.77
0	1	11	3	0.398	0.13
0	0	34	5	0.235	0.047

... ...

Key to Iteration Types for Tables 4.2 and 4.3.

1	1	Steepest descent.
1	0	Conjugate gradients step, first step of series was steepest descent.
0	1	Restart step, arbitrary initial direction. (Beale Restart).
0	0	Conjugate gradients step, first step of series has arbitrary direction.

TABLE 4.3 Order of Occurrence of Various Iteration Types for Scaled Problem Solved by CGRADS.

Iteration Number		1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	x	x											
1	0			x	x	x	x	x	x					
0	1									x				x
0	0										x	x	x	
Iteration Number		14	15	16	17	18	19	20	21	22	23	24	25	26
1	1													
1	0													
0	1	x	x	x	x		x	x	x			x		
0	0					x				x	x		x	x
Iteration Number		27	28	29	30	31	32	33	34	35	36	37	38	39
1	1													
1	0													
0	1			x										x
0	0	x	x		x	x	x	x	x	x	x	x	x	
Iteration Number		40	41	42	43	44	45	46	47	48	49	50	51	
1	1			x				x			x		x	
1	0				x	x	x		x	x		x		
0	1		x											
0	0	x												

TABLE 4.4

Problem: Scaled Control Constraints, Deterministic Hydro-Thermal.

Algorithm: Steepest Descent (XLS =.1).

Also see Figure 4.10

Penalty Iteration	Number of Iterations	Av. Function Evals. Per Iteration	Av. %age Function Change Per Iteration	Av. %age Function Change Per Iteration
2	6	6	1.055	0.18
3	272	4	0.255	0.064
4	150	3	0.409	0.136

Problem: Unscaled Control Constraints, Deterministic Hydro Thermal.

Algorithm: Steepest Descent (XLS = 1).

Also see Figure 4.11

2	457	4	0.0929	0.023
3	1137	5	0.0890	0.018
4	307	5	0.0744	0.015

TABLE 4.5

Problem: "Gaussian + Impulse" Stochastic One Reservoir Model, Two Cycles, (54 Control Variables).

Algorithm: CGRADS (XLS = 0.2, RSC = 0.5).

Iteration Type		Number of Iterations	Av. Function Evals. Per Iteration	Av. %age Function Change Per Iteration	Av. %age Function Change Per Function Evaluation
2nd Penalty Iteration:					
1	1	3	9	10.74	1.19
1	0	2	11	0.56	0.051
0	1	3	4	0.022	0.005
0	0	24	8	0.016	0.002
3rd Penalty Iteration:					
1	1	3	9	9.53	1.06
1	0	9	11	0.50	0.046
0	1	13	6	0.020	0.003
0	0	29	7	0.007	0.001

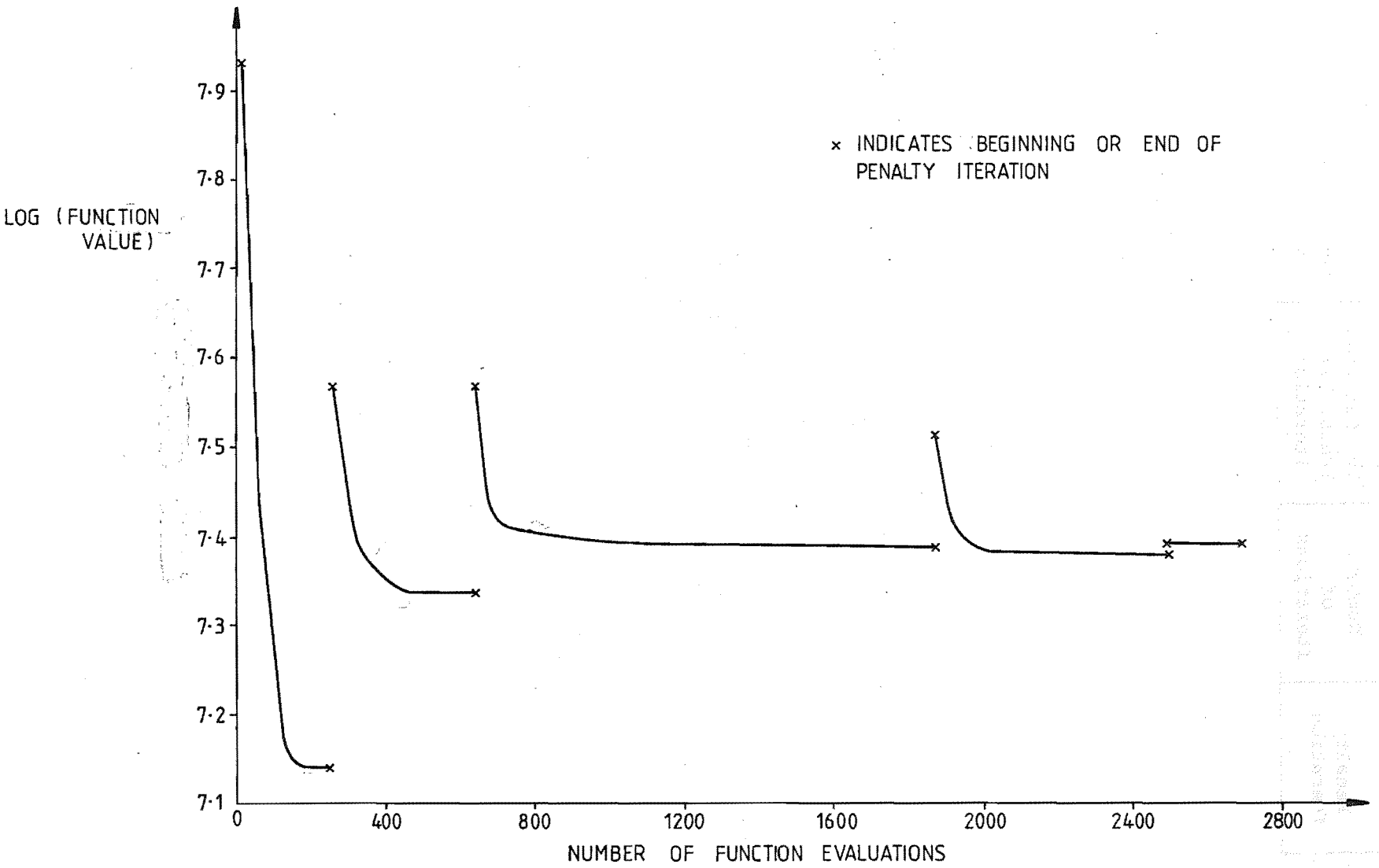


FIGURE 4.10 Reduction in function value, scaled hydro-thermal problem, solved by Steepest Descent (XLS=0.1)

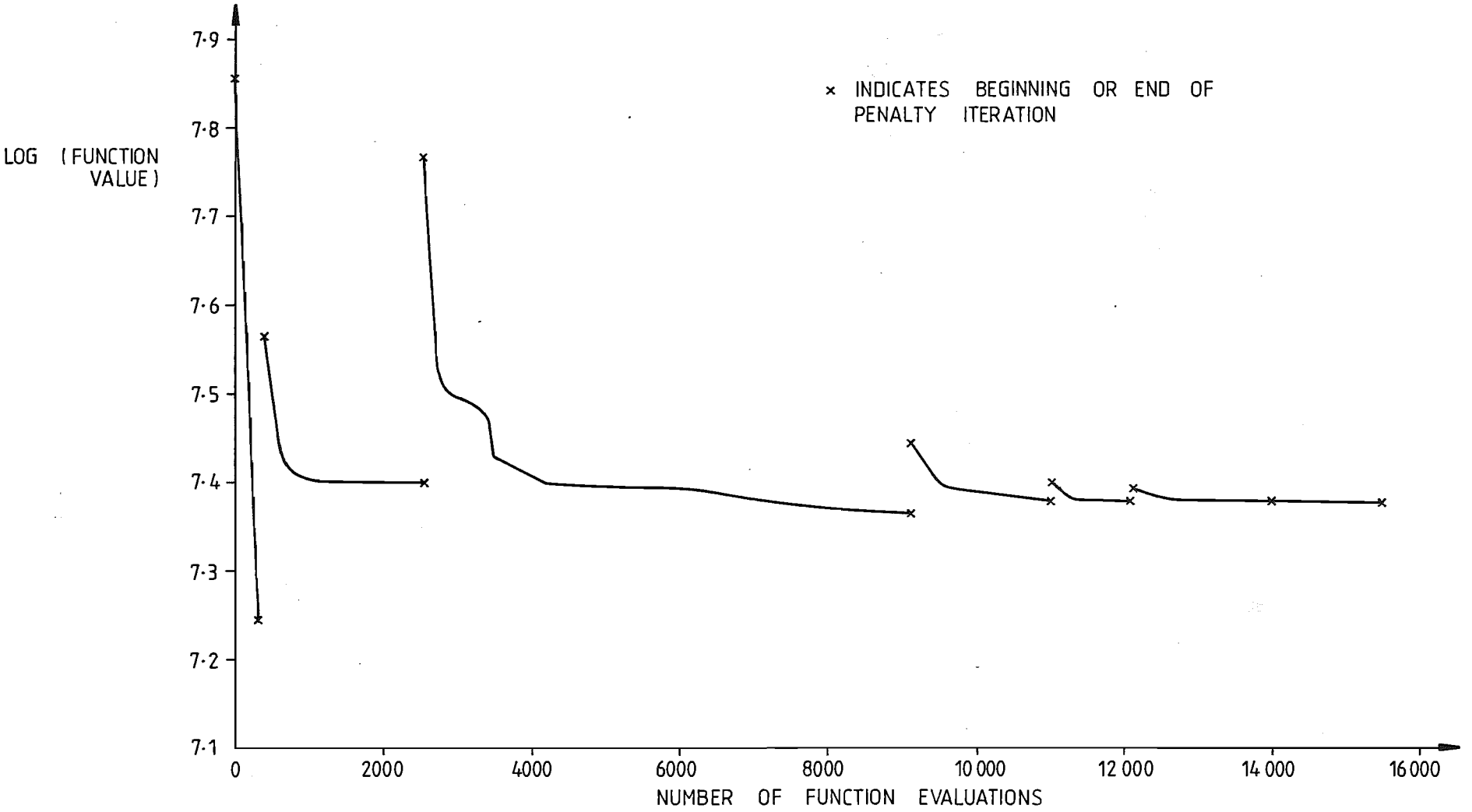


FIGURE 4.11 Reduction in function value, unscaled hydro-thermal problem, solved by Steepest Descent (XLS=1.0)

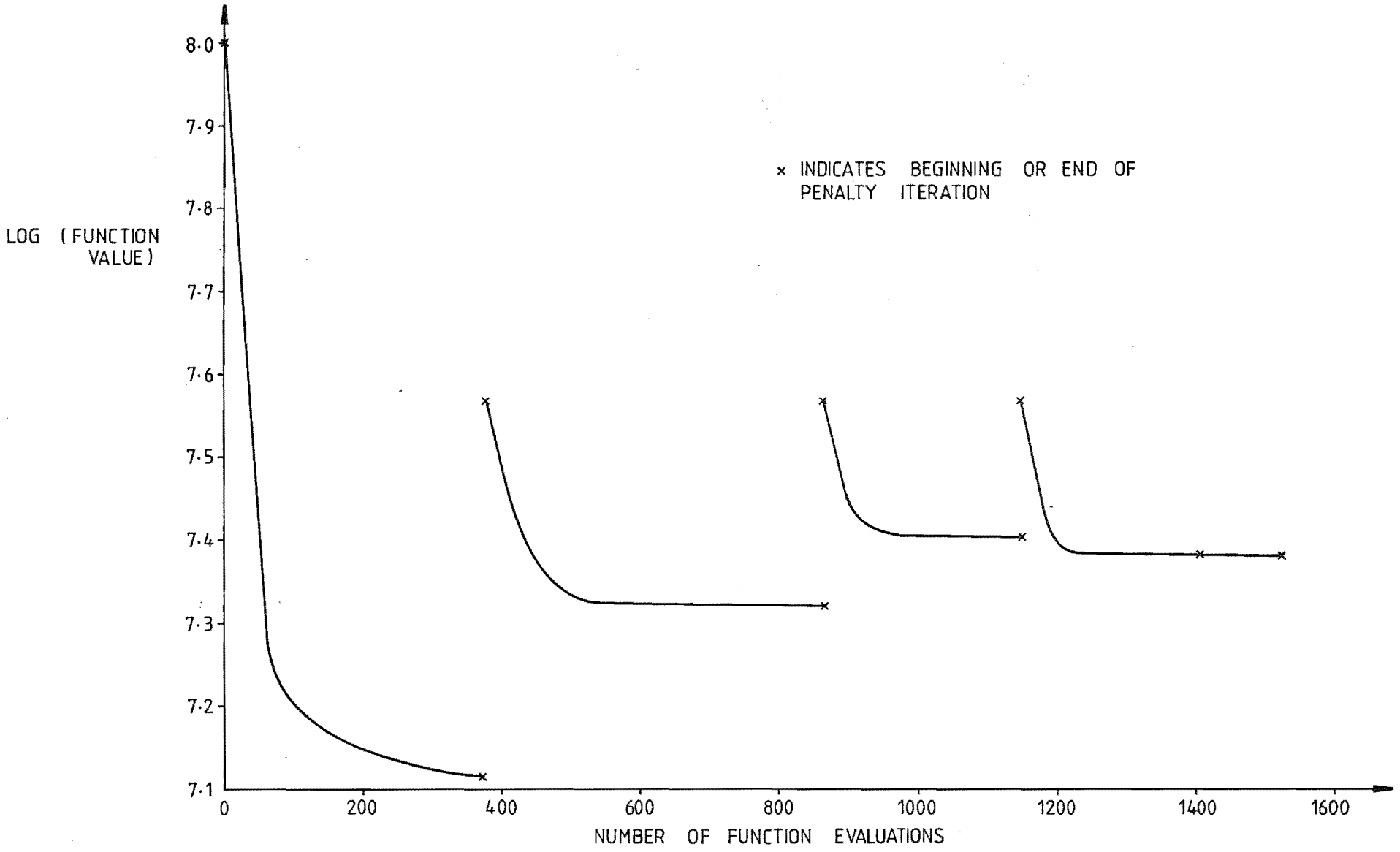


FIGURE 4.12 Reduction in function value, scaled hydro-thermal problem, solved by CGRADS (XLS=0.2, RSC=0.5)

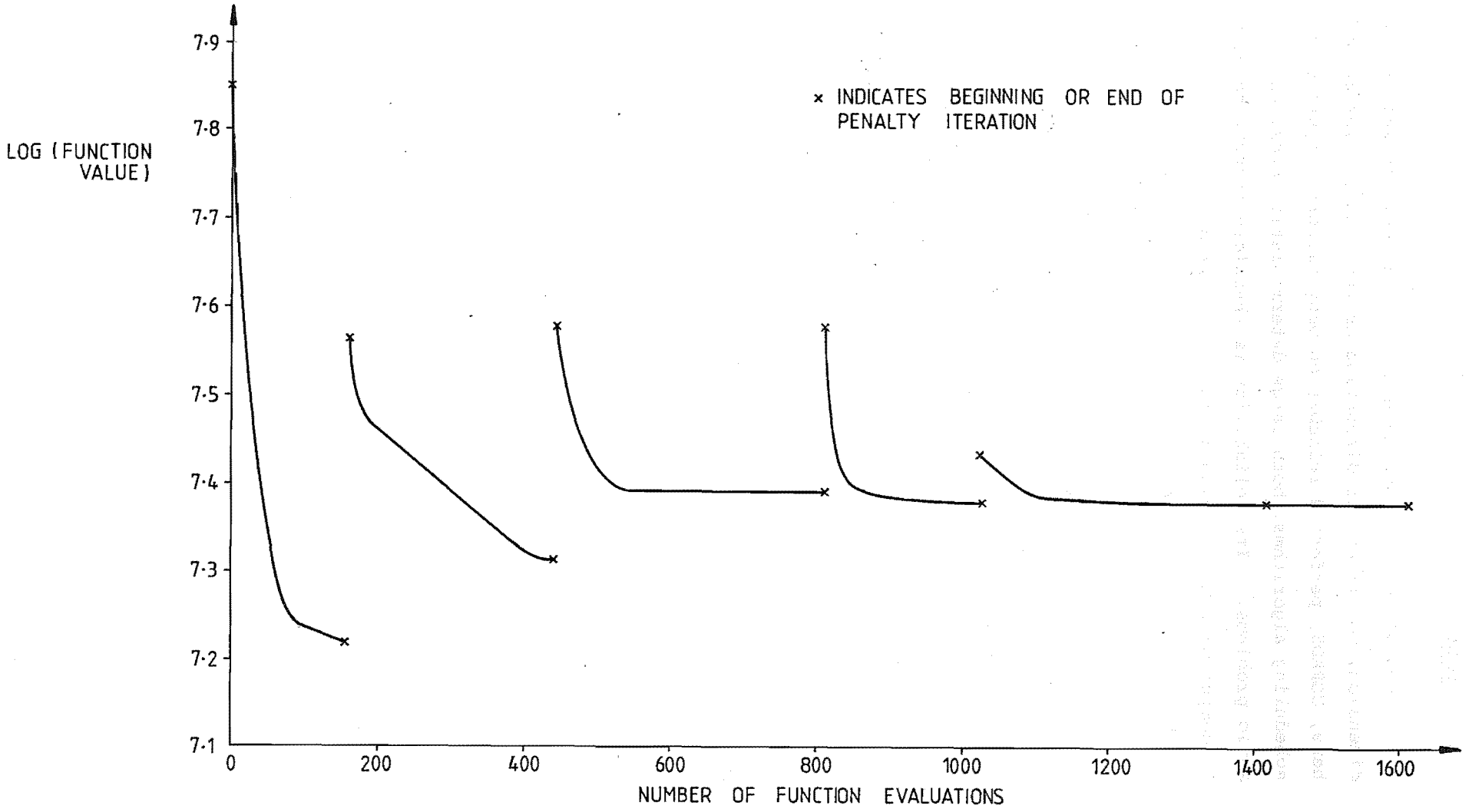


FIGURE 4.13 Reduction in function value, unscaled hydro-thermal problem, solved by CGRADS (XLS=0.2, RSC=0.5)

4.6 CONCLUSIONS

Conjugate gradients methods are particularly well suited to high dimension, non linear or discontinuous functions. The method developed here, CGRADS, performed reliably on many different types of hydro-thermal scheduling algorithms, both large deterministic forms, and smaller stochastic problems. Its reliability is therefore thought to be very good. As computing power increases, (and its cost decreases) reliability is increasing in importance over efficiency of computation.

The basic Fletcher and Reeves method, implemented by the IBM subroutine FMCG is effective on quadratic problems, but can not handle the non linear problems solved by CGRADS.

A steepest descent method took almost nine times as much computer time as CGRADS for a poorly scaled problem. This is a very significant difference in process time and justifies the added complexity of CGRADS.

Stochastic problems were solved using finite differences to determine gradients, as opposed to the analytic method used normally. Provided the increments made to control variables for finding gradients were less than the desired convergence, satisfactory performance was obtained.

CHAPTER 5

MODELLING THE NEW ZEALAND POWER SYSTEM

5.1 INTRODUCTION

In Chapter 3, a generalised model for the long term scheduling of a power system was given, with approximations especially for the New Zealand power system. The modelling of this system is described in detail in this chapter and the approximations are justified. Inclusion of the D.C. line linking the two islands extends the model of Chapter 3. The model has been developed to a state where it has been tested almost to the greatest extent possible, unless trials were to be done in conjunction with a power system operating organisation. Solutions, and details of algorithm performance, can be found in Chapter 6.

The data on which the model is based are given in Tables 5.2 to 5.6 - six thermal stations and eight hydro reservoirs are involved. The configuration of the model is valid for the N.Z.E. financial years ending in 1981 and 1982. Prior to 1981, Ohau A was not commissioned and for later than 1982 some output from Huntly thermal station needs to be considered.

To economise on computer time, development has been with a four week time interval, although one week would be more suitable for an application. 13 time intervals make up the 52 week optimisation horizon, beginning at 1st April. This is the beginning of the N.Z.E. financial year, and all lakes in the South Island are usually full at this date, ready for the winter. This makes choosing initial and final reservoir levels easy.

Each of the eight river systems and six thermal stations of the mathematical model to be developed in this chapter is given an index number for identifying them. They are:

TABLE 5.1 List of Reservoirs and Thermal Stations Appearing in the Model

Index	Name	Includes
1	Taupo	Tokaanu, all Waikato River Stations, Wairakei geothermal.
2	Waikaremoana	Kaitawa, Piripaua, Tuai, all fed by Lake Waikaremoana.
3	Cobb	
4	Coleridge	Highbank
5	Tekapo	Tekapo A, Tekapo B.
6	Manapouri	Te Anau Storage also.
7	Hawea	Generation at Roxburgh only.
8	Pukaki	Ohau A, Benmore, Aviemore, Waitaki, D.C. link terminal.
9	"Haywards"	D.C. link terminal only.
10	New Plymouth	
11	Stratford	
12	Meremere	
13	Marsden	
14	Whirinaki	
15	Otahuhu	
16	"Shortage"	

Referring to the general description of the system of section 3.3.1:

(i) m=8 reservoirs, 6 of which are hydraulically independent.

(ii) q=5, this reservoir discharges into reservoir 8 (not q+1).

(iii) n=6, number of thermal stations, increased to n=8 to include the North Island D.C. link terminal and shortages.

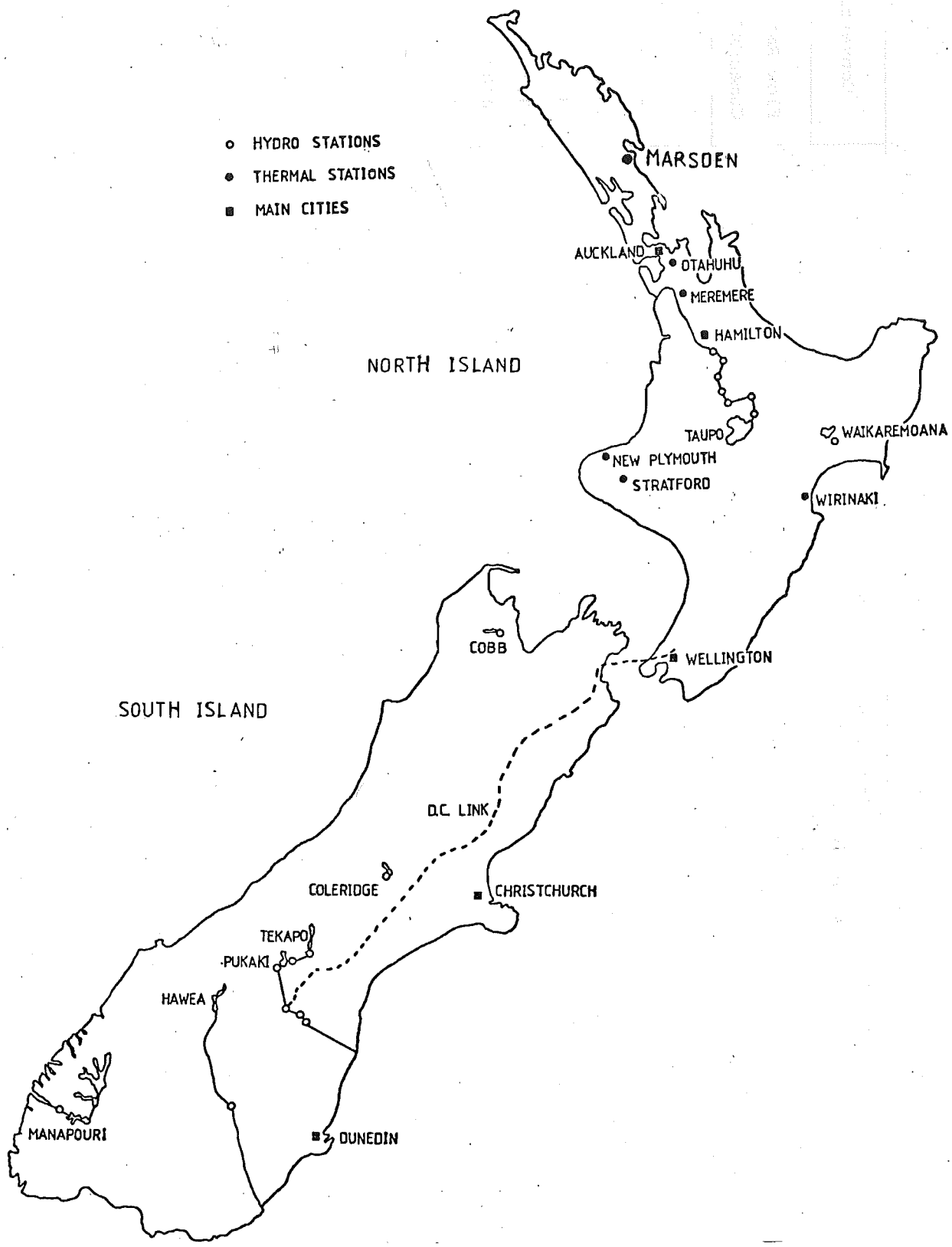


FIGURE 5.1 New Zealand map showing locations of lakes, thermal stations and main cities.

TABLE 5.2 North Island Hydro Stations - Physical Data

			Variation in Level		Lake Volume			Conversion Factors	
	Rating (MW)	Static Head (Metres)	(Metres)	(%)	(CMD)	(GWh)	Hours at Full Load to Empty Lake	Flow Rate Cumec/MW	Lake Level GWh/Metre
Taupo System									
Tokaanu	200	208	0.60	0.29	109.7	4.81	24.0	0.548	8.01
Lake Taupo	-	-	1.25	-	884.2	512.15	1061	0.41	409.72
Aratiatia	90	34	1.20	3.53	7.2	0.05	0.6	3.730	-
Ohakuri	112	35	0.45	1.29	61.2	0.41	3.7	3.596	-
Atiamuri	84	25	1.25	5.00	30.7	0.15	1.8	4.892	-
Whakamaru	100	100	0.20	0.20	14.8	0.11	1.1	3.170	-
Maraetai	360	61	0.60	0.98	33.8	0.40	0.1	2.009	-
Waipapa	51	51	0.60	1.18	11.0	0.04	0.8	7.003	-
Arapuni	158	53	0.60	1.13	62.7	0.67	4.2	2.267	-
Karapiro	90	30	1.85	6.17	156.8	1.00	11.1	3.767	-
Waikaremoana System									
Lake Waikaremoana	-	-	3.0	2.22	185.6	152.3	1350	0.29	152.3
Kaitawa	32	135				43.23		0.789	-
Piripaua	40	113	2.60	2.30	2.2	0.05	1.2	0.630	-
Tuai	52	206	1.40	0.68	4.41	0.10	1.9	1.062	-
Run-of-River									
Mangahao	192	273	1.52	0.56	42.24	1.89	9.8	0.538	1.53
Matahina	72	61	3.05	5.00	81.20	1.05	14.6	1.849	1.05

TABLE 5.3 South Island Hydro Stations - Physical Data

			Variation in Level		Lake Volume			Conversion Factors	
	Rating (MW)	Static Head (Metres)	(Metres)	(%)	(CMD)	(GWh)	Hours at Full Load to Empty Lake	Flow Rate Cumec/MW	Lake Level GWh/Metre
Cobb	32	594	15.24	2.57	279	29.41	919	0.227	1.93
Coleridge	34.5	149	3.95	2.65	1645	38.75	1123	1.020	9.81
Lake Tekapo	-	-	7.6	25.33	7791	283	1528	0.66	37.24
Tekapo A	25.2	30						4.306	-
Tekapo B	160	14.6	-	Canal Fed		-	-	0.780	-
Lake Pukaki	-	-	13.80	-	25399	1091.12	1385	0.56	79.07
Ohau A	264	54	-	Canal Fed		-	-	1.998	-
Benmore	540	92	0.85	0.92	5055	13.82	25.6	1.223	-
Aviemore	220	37	0.60	1.62	199	1.48	6.7	3.225	-
Waitaki	105	21	2.10	10.0	142	0.55	5.2	6.165	-
Lake Hawea	-	-	16.0	-	24814	234.0	2985	-	-
Roxburgh	320	46	1.85	4.02	120	1.14	3.56	2.55	2.44
Manapouri	600	177	1.80	1.02	4432	140.89		0.755	51.23
Te Anau	-	-	1.25	-	7537	239.60	634	-	129.51
Run-of-River									
Arnold	3.1	13	0.90	6.92	-	-	-	3.96	-
Highbank	25.2	101	0.55	0.54	-	-	-	1.20	-
Monowai	6	47	2.70	5.74	865	6.59	1098	3.15	-

TABLE 5.4 North Island Hydro Stations - Performance

	Year Ended 31st March 1981			Year Ended 31st March 1982		
	Total Generation (GWh)	Load Factor (%)	Availability (%)	Total Generation (GWh)	Load Factor (%)	Availability (%)
Taupo System						
Tokoanu	864.1	48.17	69.46	841.2	47.45	83.10
Aratiatia	398.2	52.91	92.23	386.3	51.76	93.94
Ohakuri	471.3	46.03	93.04	456.7	45.37	94.14
Atiamuri	346.7	48.26	92.04	341.5	48.12	91.51
Whakamaru	574.0	63.01	95.01	573.8	63.41	94.15
Maraetai	981.4	37.22	94.18	967.8	36.58	93.01
Waipapa	294.1	59.96	93.54	280.5	57.49	86.66
Arapuni	959.9	70.83	88.27	945.2	70.43	90.87
Karapiro	589.2	71.17	88.89	524.3	61.89	69.46
Waikaremoana System						
Kaitawa			92.76			96.46
Piripaua	593.7	49.54	92.56	610.37	51.42	87.93
Tuai			94.38			93.74
Run-of-River						
Mangahao	100.4	55.08	83.96	90.8	50.08	81.82
Matahina	291.5	45.03	96.11	327.5	50.72	95.60

TABLE 5.5 South Island Hydro Stations - Performance

	Year Ended 31st March 1981			Year Ended 31st March 1982		
	Total Generation (GWh)	Load Factor (%)	Availability (%)	Total Generation (GWh)	Load Factor (%)	Availability (%)
Cobb	176.4	59.41	91.23	150.4	50.64	61.92
Coleridge	223.9	67.61	87.19	241.2	71.71	83.42
Lake Tekapo						
Tekapo A	116.3	47.94	84.49	134.1	53.51	93.50
Tekapo B	711.5	56.06	72.71	751.2	54.20	77.75
Lake Pukaki						
Ohau A	802.3	35.09	78.31	1088.1	47.17	86.03
Benmore	2453.1	49.60	95.72	2451.3	48.54	95.28
Aviemore	990.6	47.64	95.05	979.3	47.17	91.94
Waitaki	532.8	61.87	89.12	533.7	55.14	93.96
Lake Hawea						
Roxburgh	1444.2	49.69	94.33	1626.9	56.59	95.84
Manapouri	4302.7	83.15	84.80	4373.7	83.35	81.82
Run-of-River						
Arnold	25.9	86.91	90.16	23.6	81.78	82.30
Highbank	67.6	36.70	80.54	92.9	42.62	88.96
Monowai	40.1	72.68	-	36.6	65.34	-

TABLE 5.6 Thermal Stations Data.

				AVERAGE PERFORMANCE			
				Year Ended 31st March 1981		Year Ended 31st March 1982	
	Rating (MW)	Fuel Type	Fuel Cost $/MWh	Load Factor %	Availability %	Load Factor %	Availability %
Marsden	240	Heavy Oil	83.4	0	66.73	0	76.04
Otahuhu Mcs 1 & 2 Mcs 3 - 6	180	Light Oil	119.1(a) 143.0(b)	0.07	51.44	1.07	59.36
Meremere	210	Coal	22.0	26.18	83.67	21.17	46.39
Whirinaki	108	Light Oil	111.8	0.18	94.84	0.15	90.64
Stratford	208	Gas	16.90	6.76	79.85	14.77	81.51
New Plymouth	600	Gas	13.50	39.24	48.76	40.93	58.51
*Huntly	1000	Coal	14.70				
Geothermal:							
Wairakei	192		0	91.42	79.76	86.61	77.16

* Under Construction

5.2 DEMAND

Historical data has been used, although in practice predicted figures are available. N.Z.E. produce forecasts for each island, of weekly GWh demands and peak MW requirements for about two years ahead. Annual predictions for 30 years ahead are produced. The load duration curve for the model is approximated by three segments (see figure 3.1) - peak (32 hours), intermediate (320 hours), and off-peak (320 hours). Test solutions of the problem using various numbers of segments would be useful to determine by trial and error how many segments are necessary. Three were used here to keep computation down.

The calculation of D.C. loadflow loss coefficients requires data for each bus, and for the three load segments for all thirteen time intervals. N.Z.E. were able to provide half hourly demands for each island in the form of weekly load duration curves. Only total weekly demands were available for each of the 74 N.Z.E. customers (power boards and large industrial users). The weekly data were first converted to four weekly figures. To apportion demand data between the three load segments, each power board's load was assumed to follow the same pattern as the total load for that island, for the particular four week period.

Once the loss coefficients have been calculated (before the actual optimisation begins) only total North and South Island load data are required. This consists of only 78 numbers, as opposed to 2886 (i.e. 39 x 74) for the loss coefficient calculations

5.3 HYDRO STATIONS

5.3.1 Reservoir Modelling Principles

Reservoir inflows are expressed in MWh terms as all hydrostations are considered to have a linear water to energy conversion factor, equation (3.2). For the Waitaki and Benmore examples of Figures 3.2 and 3.3, all machines in use are equally loaded. The lowest point of each curve of the set corresponds to the most efficient operating point for the given number of machines in use. These points rise by only 1.2% in going from one machine to all six at Benmore, but not at all at Waitaki. This makes the linear conversion factor approximation seem reasonable, in this case.

Head variations also affect the electrical equivalent value of water. Referring to Tables 5.2 and 5.3, the three largest percentage head variations are Tekepo A with 25.33% (25.2 MW installed capacity), Waitaki with 10.0% (105 MW installed and Karapiro with 6.17% (90 MW installed). All other stations have less than 5% head variation, except the tiny 3.1 MW Arnold station. The whole NZE system has 3,763 MW of installed hydro capacity, by comparison. Only 5.85% of installed hydro capacity is affected by heads that vary more than 5%. As a result the great increase in complexity caused by modelling head effects is not justified. Only one station with

significant variation is fed directly by a large reservoir, the level of which is actually known. Head variations at the other stations referred to above are not known.

Approximately 50 years of inflow data is available through TIDEDA, a computerised flow data storage scheme provided by the Ministry of Works and Development. However it was more convenient to use data prepared for the linear programming model of Boshier and Lermit (1977), covering the period 1934-1975. Some of this data has been synthesised from other inflows which were thought to be closely correlated. This was necessary due to gaps in data caused by equipment failures, floods, etc. Hence it may not be reliable for statistical studies, but the last ten years data (1972-1982) is apparently of a higher quality.

The average flows for the five years 1968 to 1973 have been used for the model. An average was used to minimise the effect of any especially extreme conditions that might be present if one particular year's flows had been used.

5.3.2 Run-of-River Stations

Data for the five run-of-river stations - Mangahao, Matahina, Arnold, Highbank and Monowai - is given in Tables 5.2 and 5.3. Monowai has a lake large enough to store water for 1,098 hours operation at full load, but its 6 MW installed capacity is insignificant compared with that of the other large reservoirs. These five stations are modelled as predetermined negative loads.

Wairakei geothermal station is treated like a run-of-river station as its output has zero incremental cost. Hence it will be used whenever it is available.

Only 3% of total hydro generation is from run-of-river stations, so they are not of great importance.

5.3.3 The Waikato System

This system shown in Figure 5.2 involves nine generating stations. Tokaanu discharges into Lake Taupo, with the diversions feeding it increasing Taupo's total inflows above their natural levels. The storage in Lake Roto Aira, above Tokaanu, is sufficient for only 24 hours full load operation, so its effect is neglected and Tokaanu's output simply added to the Waikato tributary figure.

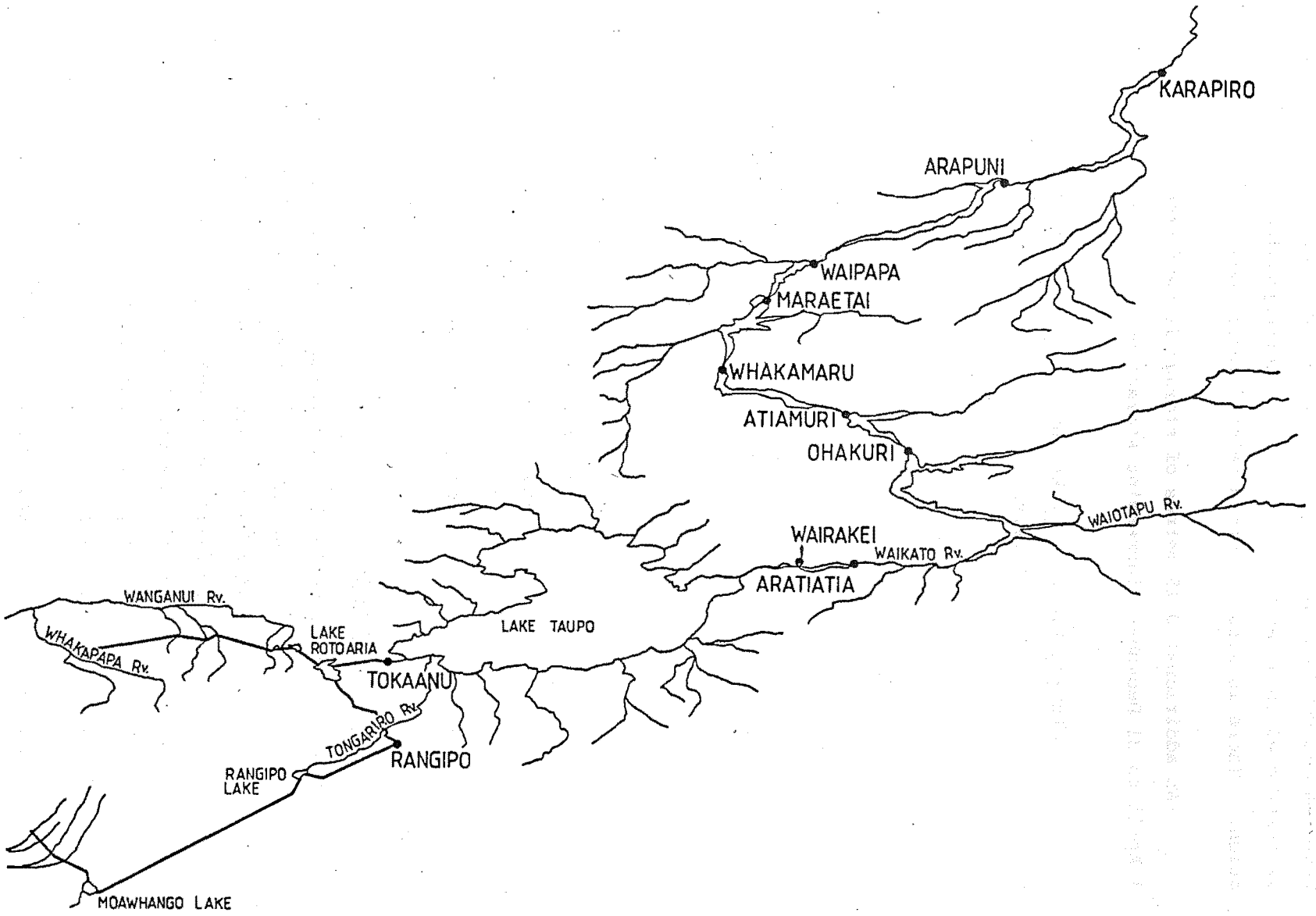


FIGURE 5.2 Waikato System Map

The eight stations on the Waikato River fed by discharges from Lake Taupo have a combined installed capacity of 1045 MW. Flow from the lake is restricted to a maximum of 200 cumec, corresponding to 488 MW of generation. There is also a minimum release requirement of 30 cumec or 73 MW.

An additional 0.15 metres of storage is permitted in the lake from 1 April to 31 December, increasing storage by 12%.

TABLE 5.7 Storage Limits on Lake Taupo, Applying at the Beginning of Various Time Intervals

Time Interval No.	1	2	3	4	5	6	7
Storage (GWh)	573.6	522	522	522	522	522	522
Time Interval No.	8	9	10	11	12	13	14
Storage (GWh)	522	522	522	522	573.6	573.6	573.6

Difference equation (3.3) and equation (3.4) describe Taupo operation with the constraints:

Minimum release rate = 73 MW
Maximum release rate = 488 MW
Minimum storage = 0
Maximum storage - see Table 5.7

5.3.4 Waikaremoana

Three stations are fed by Lake Waikaremoana, being Kaitawa, Tuai, and Piripaua in that order. The lake suffers from a strange leakage problem, with about 5 cumecs leaking past Kaitawa - a constant 6.3 MW generation potential loss - and reappearing above Tuai. Approximately a half cumec leaks past Tuai. Leakage flows alone (i.e. controllable release is zero) give a generation of about 13.5 MW. Some efforts have been made to eliminate the leakage (McPike, 1981).

Lake inflows are measured by recording flows past Tuai. The water to electricity conversion factors of Table 5.2 assume that 76.7% of water passes through Kaitawa, and 95.6% through Tuai. A further oddity of this lake is that the spillway gates are apparently inoperable. So spilling

only occurs when the lake actually overflows across the Wairoa-Rotorua roadway.

Equations (3.3) and (3.4) apply to describe this reservoir, but tributary flows are zero. Constraints are:

Maximum release rate	=	124 MW
Minimum release rate	=	13.5 MW (leakage)
Maximum storage	=	152.7 GWh
Minimum storage	=	0

5.3.5 The Waitaki System

The Waitaki river is fed by two controlled lakes - Tekapo and Pukaki. Major uncontrolled inflows come from Lake Ohau and the Ahuriri river. Figure 5.3 shows the locations of the components, although Ohau B and Ohau C stations are still under construction. They have not been modelled. The model is illustrated by Figure 5.4. Tekapo and Pukaki are modelled separately as the flow constraint of the Tekapo-Pukaki canal might require spill from Tekapo with very high inflows. Tekapo spill goes down the old Tekapo River bed, into Lake Benmore, but is modelled as being lost completely as a simplification. Likewise Pukaki spill is down the Pukaki River bed into Lake Benmore, but is also modelled as a complete loss, i.e. it is assumed to be spilt at Benmore, Aviemore, and Waitaki as well.

Lake Benmore is the largest man-made lake (i.e. one contained only by a man-made dam) in the NZE system and might need to be modelled as controlled storage if a one week rather than a four week time step were used.

Tekapo storage limits are higher in winter as shown in Table 5.8. The GWh storage figures refer to generation potential at Tekapo A and B only, as the water then flows into Pukaki where its use can again be controlled.

TABLE 5.8 Tekapo Storage Limits. Note: T=1 corresponds to 1st April.

Time Interval No.	1	2	3	4	5	6	7
Storage (GWh)	294.9	307.2	319.5	331.8	331.8	302.7	294.9
Time Interval No.	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Storage (GWh)	283	283	283	283	283	283	294.9

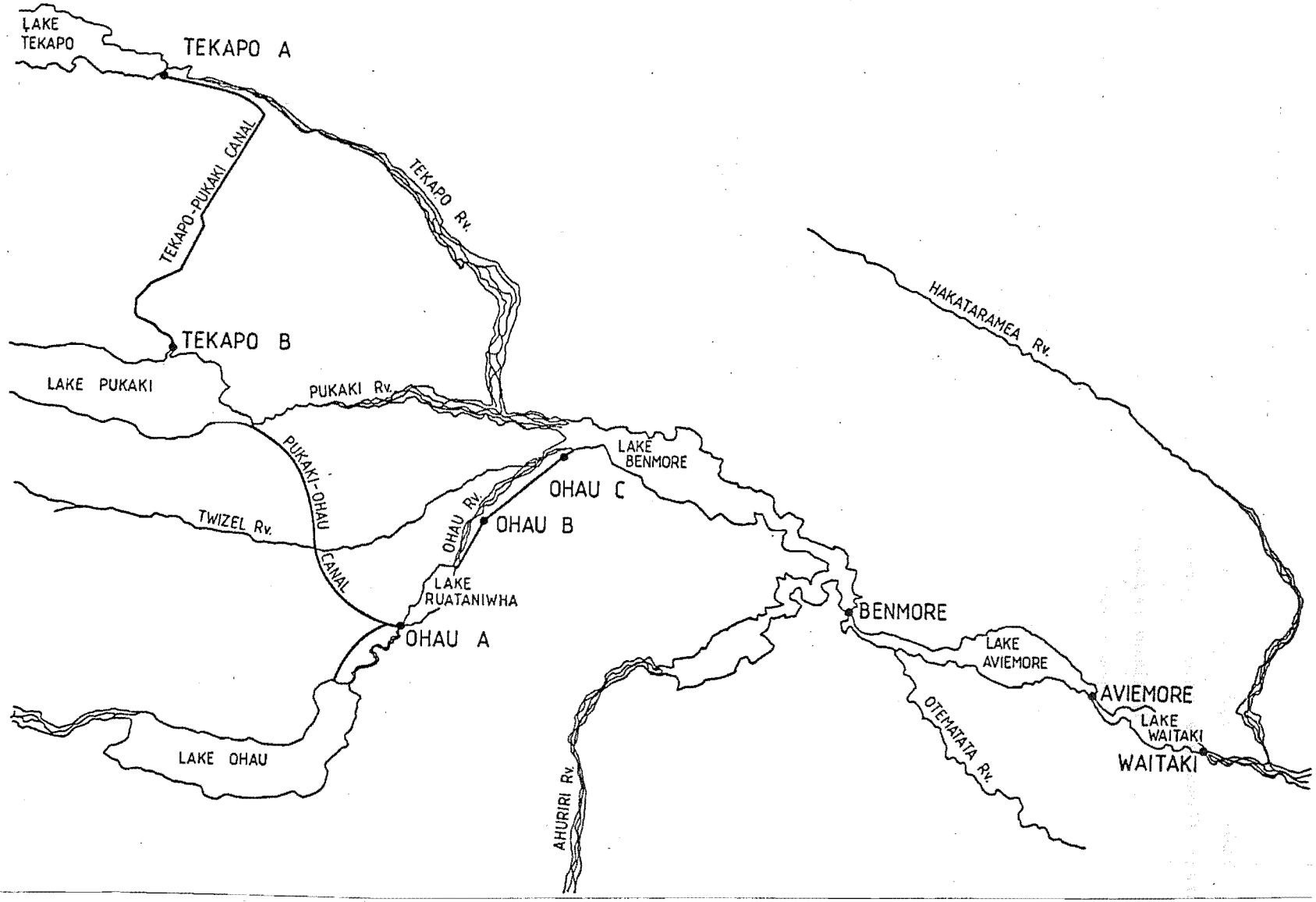


FIGURE 5.3 Waitaki System Map

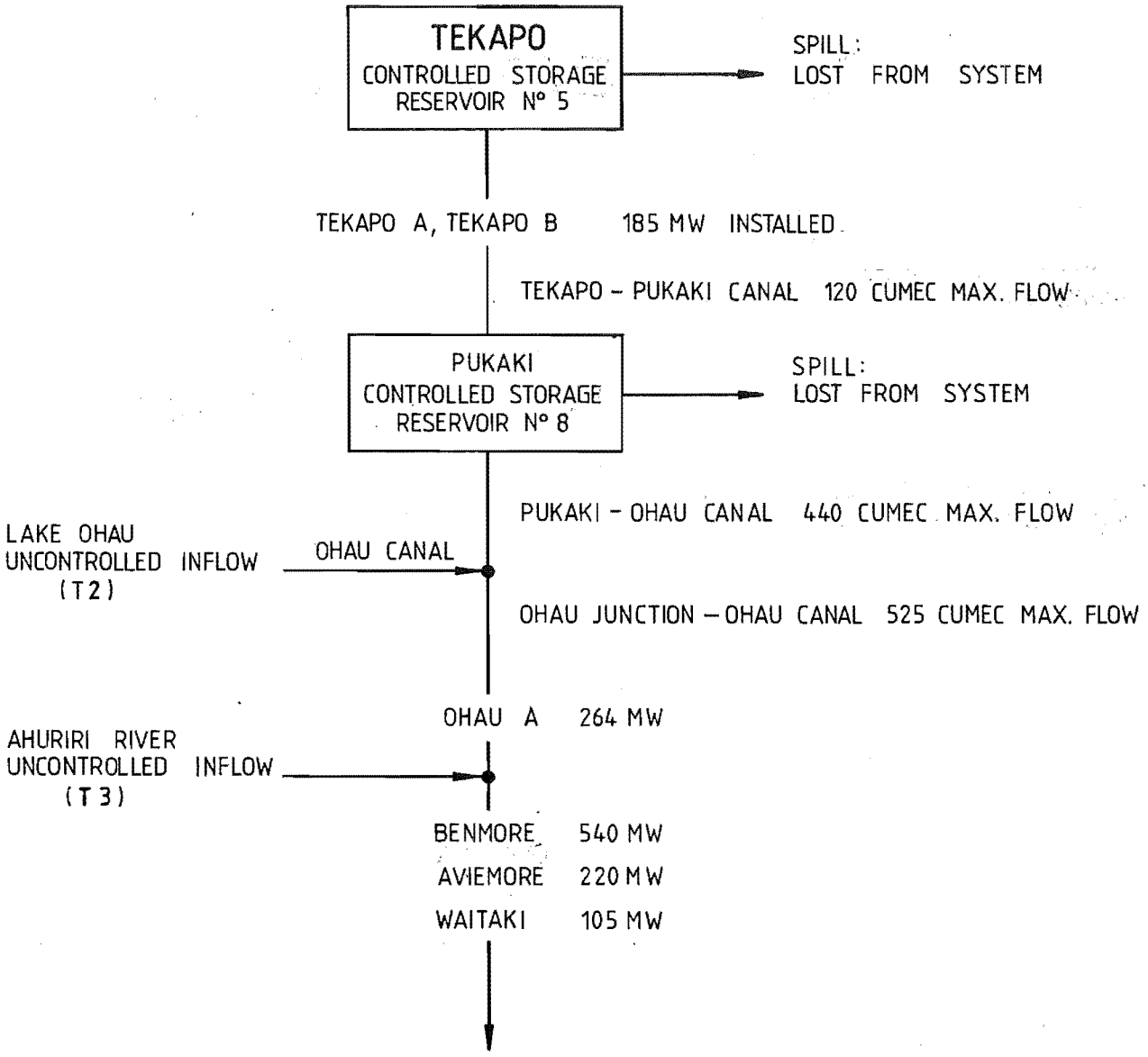


FIGURE 5.4 Model for Waitaki System

The canal carrying water from Pukaki has a 440 cumec capacity, increased to 525 cumec from the junction with the Ohau canal. This 525 cumec flow is sufficient to run Ohau A at full load of 263 MW, or a total of 940 MW if the same flow rate also applies at Benmore, Aviemore, and Waitaki. Hence the maximum discharge from Lake Pukaki is set to 525 cumec minus Ohau tributary flow, or 440 cumec, whichever is the smaller.

The Ahuriri river flows directly into Lake Benmore. It is assumed that this water can always be fully utilised. The combined installed capacity of the three lower Waitaki stations is 865 MW, but the 525 cumec continuous flow down the canal to Ohau A would give only 677 MW. Hence there is 188 MW spare installed capacity available for tributary flows, and peaking within the three load classes.

Tekapo, the top lake of the cascaded pair, is described by equations (3.3) and (3.4), with zero tributaries. Pukaki has (3.12) and (3.4) applying, but with two sets of tributary flows - Ohau and Ahuriri. The conversion factor $k_{8/5} = 1.18$, for equation (3.12).

Constraints on the two reservoirs are:

	Tekapo	Pukaki
Max. release	182 MW	$\text{Min}(788, 940 - T_2(t) \frac{h_j}{672})$ MW
where $T_2(t)$ = Ohau Tributary Flow		
Min. release	0	0
Max. storage	as per Table 5.8	1091 GWh
Min. storage	0	0

5.3.6 Other South Island Reservoirs

Cobb and Coleridge having only 32 MW and 34.5 MW respectively, installed capacity could possibly be omitted from the model. However their inclusion is useful for testing the conjugate gradients algorithm. A given percentage change in generation at Cobb has much less effect on overall costs than the same percentage change at Manapouri. This can give rise to the sort of function contours that a steepest descent method has difficulty handling, so it helps test the effectiveness of CGRADS.

Manapouri has installed seven machines of 100 MW each, but only six can be used at once without flooding the station due to excessive tailwater elevation. So capacity here is set to 600 MW, with no reduction for outages.

With the spare machine 600 MW should be available almost always. The storage of both Lakes Manapouri and Te Anau are available. The levels of these lakes can deviate from the normal operating range but only provided the rules laid down by a special act of Parliament are adhered to. These specify the duration and frequency of the deviations. Only the normal operating range has been modelled.

Roxburgh station, on the Clutha River, has uncontrolled flows from Lakes Wakatipu and Wanaka, and controlled storage from Lake Hawea. The station has 320 MW installed but the maximum outflow from Hawea of 200 cumec gives only 78 MW generation at Roxburgh, and the minimum of 6 cumec gives 2.35 MW. The maximum outflow actually drops to 135 cumec at the minimum operating level, but the higher limit has been used for the model. The bottom 37% of storage requires Ministerial approval before it can be drawn on, so such low operating levels should not be common anyway. A further restriction is that Roxburgh generation must not fall below 60 MW under normal circumstances. This is to ensure that water users further down river have adequate supplied.

Tributary flows are present for Hawea. Constraints for these reservoirs are:

	Maximum Release Rate (MW)	Minimum Release Rate (MW)	Maximum Storage (GWh)	Minimum Storage (GWh)
Cobb	32	0	29.4	0
Coleridge	34.5	0	38.7	0
Manapouri	600	0	380.5	0
Hawea	78.5	2.35	234	0

Further constraints on Hawea:

Maximum generation = 320 MW

Minimum generation = 60 MW

5.4 THERMAL STATIONS

The assumption of constant incremental fuel cost is used, as mentioned in section 3.3.4. This is because all thermal power station turbines reach maximum efficiency at full load, except at Meremere. Here the maximum occurs at 84% of full load.

A seventh station representing any shortfall in generation has an incremental cost of 200 $/MWh, whereas the most expensive real station costs 120 $/MWh.

The piecewise linear fuel cost function can give contributions to the first derivative which take only eight different values. Hence not all values are possible for the Hamiltonian's derivative. To assist convergence of the conjugate gradient algorithm, a linear transition in incremental costs from one cost to the next was introduced. There is now a constant rate of change of incremental cost over the lowest and highest 12.5 MW output of each thermal station, permitting all incremental costs from 0 to 120 $/MWh to be found, for some generation level. Generation costs must be quadratic over these transitions but total costs are distorted only during the transition, being accurate elsewhere.

Table 5.6 gives the (now outdated) fuel costs, station installed capacities, and some average performance data for the thermal stations.

5.5 TRANSMISSION LOSSES, AND THE D.C. LINK

Transmission losses in practice average about 6% of energy generated. D.C. link losses are set to a constant 10% of the energy sent, for the model, which is a rather pessimistic value.

Separate sets of loss coefficients must be calculated for the North and South Island systems, as the D.C. link must be simulated by a power station at each end. This actually reduces the number of coefficients to be calculated, i.e. instead of E_{vv} being a 13 x 13 matrix, two matrices are required - one 4 x 4 the other 8 x 8. As each island has its own slack bus which does not appear in that island's loss calculation, computation is further reduced.

The southern terminal of the D.C. line is combined with the outputs from Pukaki and Tekapo. These two lakes are considered to feed one bus for loss calculation purposes. An extra generating station, "Haywards", is added to the North Island system, as the other D.C. line terminal. Taupo and Waikaremoana also feed only one bus each as individual station outputs are not known. All three Waikaremoana stations feed transmission lines from Tuai, and Whakamaru is a major switching point fed by several Waikato stations. So the approximations seem reasonable.

In the North Island, the most expensive thermal station in use is the slack bus for loss calculations. "Haywards" is used if no thermal stations are operating. In the South Island the Tekapo-Pukaki bus is always slack as any change in losses or generation is compensated for by a change in D.C. link power, fed from this bus.

A total of 210 lines and 81 busses are modelled for loss coefficient calculations. Paralleled lines are treated as one, and some minor lines have been omitted. The total losses are:

$$L = L_{SI} + L_{NI} + L_{DC} \tag{5.1}$$

$$= \sum_{v(SI)} \sum_{N(SI)} G_{v(SI)} ECC_{v,N(SI)} G_{N(SI)} + \sum_{v(SI)} EC0_{v(SI)} + E00_{(SI)}$$

$$+ \sum_{v(NI)} \sum_{N(NS)} G_{v(NI)} ECC_{v,N(NI)} G_{N(NI)} + \sum_{v(NI)} EC0_{v(NI)} + E00_{(NI)}$$

$$+ (1-\varepsilon_{DC}) G_8 \cdot (1-\sigma(G_8)) + \frac{1-\varepsilon_{DC}}{\varepsilon_{DC}} G_8 \cdot \sigma(G_8)$$

where: L_{SI} = total South Island losses

L_{NI} = total North Island losses

L_{DC} = D.C. link losses

$ECC_{(SI)}$ etc. are South Island loss coefficients

$\sum_{v(SI)}$ indicates a summation over all South Island generations, except Pukaki and Tekapo, which are always "slack".

ε_{DC} = D.C. link efficiency

$\sum_{v(NI)}$ indicates a summation over all North Island generations, including Haywards and the thermal stations, except the slack station.

The D.C. link has a maximum sending end capacity of 620 MW. This is the only transmission constraint enforced. A square law penalty is used, as for the state constraints.

5.6 CONCLUSIONS

The New Zealand system has been described in terms of the general model of Chapter 3. Chapter 6 will give details of the solutions obtained for this example. Complications have been caused by modelling a D.C. link, but the changes affect only the loss coefficient calculations seriously. No difficulty is experienced with the Upper Waitaki system of two controlled lakes, one uncontrolled lake, one major tributary flow, six hydro stations, and three canals of differing capacities.

The assumptions made in Chapter 3 which simplify the model (regarding thermal fuel costs, and linearity of water to energy conversion) have been shown to hold. The principle weakness may be in the representation of the chain of eight stations on the Waikato river as one aggregate station. No firm conclusions can be drawn here.

All constraints considered important for the present operations planning methods have been represented, yet the model is still simple enough to be solved easily.

CHAPTER 6

SOLUTIONS FOR THE NEW ZEALAND SYSTEM

6.1 INTRODUCTION

Chapter 3 described a general modelling technique for the deterministic scheduling problem, along with a solution method. This involved the use of the conjugate gradients static optimisation algorithm of Chapter 4. Application of the modelling techniques to the New Zealand system were described in Chapter 5. Now the results of this model will be given.

After a brief summary of the model, the computer outputs for two different years are discussed - one in some detail. The objective is to demonstrate that the results are reasonable. This will indicate whether any gross errors are present due to erroneous assumptions in developing the model.

The use of the scheduling algorithm as a possible aid to system development is shown by presenting some results for the New Zealand system with various levels of lake inflows, and other changes. Finally, the various constraint enforcement methods are compared. In addition to the use of penalties or either of the two transformations to enforce controls, the variation of the form of the penalty function itself is examined in detail.

6.2 SUMMARY OF PROBLEM AND SOLUTION PROCESS

Reservoirs and thermal stations modelled are listed in Table 5.1. Their features are summarised by Tables 5.2 to 5.6 and locations are shown on the map, Figure 5.1.

The model is valid for the years ended 31 March 1981 and 1982.

Unless stated otherwise, constraints are enforced by a square law penalty function for states (reservoir storage) and a scaled transformation for controls (release). Usually all penalty weights are set to 0.1×10^{-4} initially. The CGRADS hill climbing method is used throughout this chapter except for some specific tests with other methods.

The solution process begins with the calculation of loss coefficients. This uses line, bus load, and run-of-river stations generation data. Total demands for each island are then read, and the predetermined run-of-river generations subtracted. Inflow data is read next, enabling reservoir release limits to be determined, as some limits are affected by tributary flow levels.

The optimisation process commences from the specified initial storage in each reservoir, and guess for the releases (arbitrary values will do). The general procedure is shown in Figure 3.5, using the cost function, eqn (3.15) with costates and gradients as given in section 3.6.4.

After each penalty iteration has converged, constraint violations are checked. A tolerance of 2% of the maximum value of the variable concerned is allowed at the solution. Penalty weights on variables violating constraints are increased by a factor of 10, and another iteration performed, if required.

6.3 OPTIMISATION OUTPUTS FOR 1981 AND 1982

6.3.1 Introduction

The results for the year ended 31 March 1982 will be presented in detail. The objective is to show that the solutions obtained are reasonable. A comparison for the years ended 31 March 1981 and 1982 will also be made to show that the optimisation outputs and the data given in Ministry of Energy annual reports are similar.

In both bases the model uses 1968-1973 average water inflows but thermal station availabilities are taken from the published annual reports, as in Tables 5.4 to 5.6. Demand data is for the year ended 31 March 1978. Total 1981 generation was 3.6% up on 1978 so all demands have been scaled upwards by this amount. For the 1982 case, a 7% increase was made, in the same way.

TABLE 6.1 Comparison of Optimisation Output and N.Z.E. Annual Report, Year Ended 31st March 1982.

	Optimisation Output	N.Z.E. Annual Report For Y/E 31st March 1982
Fuel Cost	$39.26 Million	$40.37 Million
Thermal Generations: (GWh)		
Huntly	-	16.4
New Plymouth	1694.6	1746.3
Stratford	364.9	304.0
Meremere	95.5	262.0
Marsden	75.3	0
Whirinaki	50.0	1.5
Otahuhu	4.5	11.4
Shortfall	3.5	-
Hydro Generations: (GWh)		
Taupo	5399.1	5364.21
Waikaremoana	342.0	610.4
Cobb	156.8	150.4
Coleridge	258.1	241.2
Tekapo	947.6	885.2
Manapouri	4733.4	4373.7
Hawea	1412.2	1626.9
Pukaki	4470.2	5052.3
Excluding Run of River		
(a) Total N.I. Hydro (GWh)	5741.1	5974.6
(b) Total S.I. Hydro (GWh)	11978	12329
% N.I. Flows Spilled	0	4
% S.I. Flows Spilled	0	14
Total Power Sent North by D.C. Link (GWh)	4502.2	4884
Average D.C. Link Power (MW)	513.9	557.5
Total Generation by N.Z.E., All Stations. (GWh)	21104	22239

Note: Demands used by optimisation are 1.07 x y.e. 31st March 1978 figures.

TABLE 6.2 Comparison of Optimisation Output and N.Z.E. Annual Report, Year Ended 31st March 1981.

	Optimisation Output	N.Z.E. Annual Report For Y/E 31st March 1981
Fuel Cost	$24.938 Million	$32.072 Million
Thermal Generation (GWh)		
Huntly	0	0
New Plymouth	1066.8	1502.3
Stratford	340.2	129.1
Meremere	132.9	393.4
Marsden	20.2	0
Whirinaki	1.06	2.7
Otahuhu	0	0.52
Shortfall	0	-
Hydro Generations: (GWh)		
Taupo	5405.6	5184.8
Waikaremoana	342.1	593.6
Cobb	157.2	176.4
Coleridge	257.9	223.9
Tekapo	947.4	857.8
Manapouri	4731.0	4302.7
Hawea	1410.1	1444.2
Pukaki	4467.5	4778.7
Excluding Run of River		
(a) Total N.I. Hydro (GWh)	5747.7	5778.4
(b) Total S.I. Hydro (GWh)	11971.1	11783.7
% N.I. flows spilled	0	4
% S.I. flows spilled	0	23
Total power sent north by D.C. Link (GWh)	4731.3	4554
Average D.C. Link Power (MW)	540.10	519.9
Total Generation by N.Z.E., All stations (GWh)	20490	21539

Note: Demands used by optimisation are 1.036 x Y E 31st March 1978 figures.

TABLE 6.3 Initial Storages and Final Time Targets for Optimisations, and From N.Z.E. Annual Report Data.

Reservoir	OPTIMISATION DATA				N.Z.E. ANNUAL REPORTS					
	Specified Initial Storage		Minimum Final Storage		31st March 1980		31st March 1981		31st March 1982	
	% Full	GWh	% Full	GWh	% Full	GWh	% Full	GWh	% Full	GWh
Taupo	50.5	290	50.5	290	72.8	372.9	54.4	278.6	48.8	249.9
Waikaremoana	50.0	76.0	50.0	76.0	81.2	117.0	36.5	52.6	33.1	50.6
Cobb	99.8	29.4	90.0	26.5	-	-	-	-	-	-
Coleridge	99.9	38.7	90.1	34.9	-	-	-	-	-	-
Tekapo	98.3	290	98.3	290	99.7	632.8	112.3	694.0	101.4	626.6
Manapouri*	78.9	300	78.9	300	98.0	400.2	76.0	289.6	81.6	310.5
Hawea	98.3	230	90.0	210.6	97.9	279.0	116.6	272.8	100.0	234.0
Pukaki	91.7	1000	871.1	950	101.1	1122.2	102.6	1119.6	98.7	1077.4

* "Manapouri" includes Te Anau storage.

6.3.2 Comparison of Annual Statistics

Tables 6.1 and 6.2 compare the 1982 and 1981 NZE annual statistics with the simulating optimisations. The most significant difference is the total annual generation - over 5% in both cases between the NZE and optimisation figures. This suggests an omission in the demand data we used.

Total hydro outputs for both optimisations are almost unchanged as the same inflow sequences have been used, and no spilling occurs in either case. Actual 1982 flows were 107% of mean resulting in spilling in both islands as shown. The five year average flows still give generation totals similar to N.Z.E. figures in both cases, except for Waikaremoana. An examination of the flow data used showed that the long term average inflows were equivalent to 491.5 GWh per year. The data used, however, gives only 341.7 GWh for the year.

Thermal station fuel cost data was obtained in July 1981, costs ranging from 13.5 $/MWh up to 120 $/MWh. By October 1982, the maximum cost had risen to 200 $/MWh. For the 1981 case, the difference in generation at New Plymouth accounts for most of the higher thermal fuel bill of the N.Z.E., over the optimisation figure.

In both years N.Z.E. did not generate at Marsden, although it was available, instead using the more costly Otahuhu and Whirinaki stations. As these two have gas turbines, they are more suitable for supplying short duration peaks. The optimisation specified generation at Marsden, but it could easily be rerun with a zero availability at that station. Then, if Otahuhu and Whirinaki use was excessive, the optimisation could be repeated with Marsden available. All three stations are used as synchronous condensers for voltage support, in practice.

Huntly generated 16.4 GWh in the 1982 year, but is not modelled. It was first used in December 1981, which is not a peak load time, so it will have had little effect on total costs, (i.e. it probably only replaced low cost New Plymouth generation during time = 10 of the optimisation.

A 3.5 GWh shortfall in peak energy supply occurs in the 1982 optimisation. All generation outages have been averaged over the year, but in practice planned maintenance would avoid the winter peaks as far as possible. If New Plymouth availability was raised by 15% from its 58.5% annual average to 73.5% for time = 4 and 5 peaks, (a total of 64 hours) then the shortfall would not have occurred.

Table 6.3 shows the initial storages used for reservoirs for all optimisation runs, including the two discussed here. The maximum storage level for Tekapo is higher for a period after 31 March, which probably accounts for the 101.4% full figure. The levels used for the optimisation can be seen to be similar to those applying in practice. As no storage constraints, except the final time target levels, are active at the solution, only the difference between initial and final levels has any effect on the solution.

In conclusion, there are no gross or inexplicable differences which have been found between the optimisation and the N.Z.E. annual statistics.

6.3.3 Validity of the 1982 Simulation

The output of the 1982 year simulation is shown in Table 6.5 and Figure 6.1. A detailed discussion of results follows.

Determining whether convergence has been obtained can be checked in a number of ways. The most simple is to rerun the computer program with a smaller control variable tolerance, EPSU, in the conjugate gradients algorithm. From Table 6.4 it appears that convergence is obtained with EPSU = 5.0, as EPSU = 0.5 gives a solution only 0.115% different.

TABLE 6.4 Cost of Thermal Fuel for Various Control Convergence Tolerances 1982 Year Simulation

EPSU	Cost of Thermal Station Fuel	% of Minimum Cost Attained
50.0	$40.9174 million	103.573
5.0	$39.5058 million	100.000
0.5	$39.5511 million	100.115

A comparison of the North Island reservoir release pattern with the incremental fuel costs is the next step.

These incremental costs, Table 6.5(g), peak for the year at t=4 and 5, corresponding to July and August. Lowest costs come during January, t=11. Off peak load segment costs are always 13.5, but six of these require less than 25% of maximum output from New Plymouth. Fuel costs and the relatively small effect of losses are the two external factors influencing the schedules of Taupo and Waikaremoana.

The Taupo release pattern can be summarised as:

Thermal Fuel Cost ($/MWh)	Releases (% of Maximum)
above 16.9	above 99.42
16.9	above 98.8
above 13.5 but below 16.9	98.0 to 99.4
13.5	25.0 to 99.1

No spilling occurs, and the only state constraint active is the terminal one. Hence all costates are equal to the terminal costate (-13.62). Neglecting loss effects, the gradient $\partial H/\partial G$ which is the gradient without the control transformation factor (see eqn 3.47) is:

$$\partial H/\partial G = -(\lambda+\rho)$$

i.e. the negative of (gradient plus costate). So for the various thermal station incremental costs:

Incremental Cost, λ	13.5	16.9	22.0	83.4	111.8	120.0
Gradient, $\partial H/\partial G$	0.12	-3.3	-8.4	-69.8	-98.2	-106.4

An examination of actual values for $\partial H/\partial G$ printed out by the computer on convergence shows that loss effects contribute at most about 0.03 λ to the gradient value. This is not enough to cause overlapping of the range of gradients possible with different incremental fuel costs. The expected release pattern is therefore:

$\lambda > -\rho \Rightarrow$ maximum release

$\lambda \simeq -\rho \Rightarrow$ any release

$\lambda < -\rho \Rightarrow$ minimum release.

As the average inflow into Lake Taupo is 421 MW and the maximum release only 488 MW, there is not a great deal of scope for variation in most of the releases. Hence all releases for prices above the minimum price of 13.5 $/MWh are greater than 98% of the maximum release - effectively at the maximum. This is consistent with the gradient $\partial H/\partial G$ being -3.3, or more negative, for prices of 16.9 or greater. All prices of 13.5 $/MWh correspond to releases from 25% to 99.1% of maximum - again consistent. All the gradients here are close to zero (i.e. 0.12).

Waikaremoana presents more scope for big differences in release, for the various thermal prices. Average inflow is only 39 MW but maximum release is 111.6 MW, and the minimum 13.5 MW. As for Taupo, all costates are equal to the terminal costate, - 17.0. So neglecting losses again, $\partial H/\partial G$ takes the values:

Incremental Cost λ	13.5	16.9	22.0	83.4	111.8	120.0
Gradient $\partial H/\partial G$	3.5	0.1	-5.0	-66.4	-94.8	-103.0

The release pattern, for various thermal fuel prices is:

Thermal Fuel Cost ($/MWh)	Releases (% of Maximum)
above 22.0	above 99.84
above 16.9	above 96.5
16.9	1.63 to 95.3
13.5 to 16.9, exclusive	1.02 to 52.5
13.5	below 1.37, except for one at 34.

For all occasions on which $\partial H/\partial G$ is negative, releases are near maximum. Only small differences in gradients exist for other prices, the range being 0.1 up to 3.5. All releases except one, for which the lowest price applies are effectively at their minimum possible value. This exception is for t=11, load segment 1. Reducing this release to the minimum and increasing other releases that are with 16.9 $/MWh costs, by the same amount, would save $3,480 of the $40 million annual fuel bill (about 1 part in 10^4 saving). Prices above 13.5, up to and including 16.9, fall over almost the entire range possible, as expected.

South Island release patterns have only two requirements, other than the satisfaction of reservoir constraints. These are that South Island demand be met, or energy imported via the D.C. link at costs determined by North Island thermal generation levels. Remaining water is used such that D.C. link exports are at a maximum when thermal costs are high. All D.C. link transfers are south to north, at the solution.

For prices above 16.9 $/MWh, D.C. link power levels exceed 99.86% of maximum, except for time = 4, peak load segment. This period has only a 97% link utilisation, despite all South Island releases being at their

TABLE 6.5(a) 1982 Year Simulation - Output of Optimisation

Note: Storages in GWh. All other quantities in MW averages.

Reservoir: Taupo. Max. Release = 488.1 MW, Min. Release = 73.0 MW.

Time	Lake Releases - 3 Load Segments			Tributaries	Storages	Maximum Storage
1	483.8	483.2	291.3	142.4	290.0	573.6
2	487.5	485.7	386.6	173.1	225.4	522.0
3	488.1	486.8	483.8	198.4	177.9	522.0
4	487.8	487.7	481.8	228.9	138.8	522.0
5	487.8	487.7	479.0	228.6	148.1	522.0
6	488.1	486.0	476.9	229.8	157.8	522.0
7	487.7	483.9	463.3	229.7	170.7	522.0
8	485.3	482.8	419.7	217.8	188.5	522.0
9	484.0	485.2	341.3	208.8	201.8	522.0
10	482.4	482.8	142.0	195.0	225.3	522.0
11	478.1	479.5	122.1	173.8	291.7	522.0
12	484.9	483.6	226.3	163.9	332.2	573.6
13	483.4	481.7	280.5	156.0	322.3	573.6
14					283.1	573.6

Time:	1	2	3	4	5	6	7
Inflows:	295.9	367.8	427.3	498.7	498.1	500.9	500.7

Time:	8	9	10	11	12	13
Inflows:	422.7	451.6	419.3	369.5	346.4	327.7

... ...

Reservoir: Waikaremoana. Max. Release = 111.5 MW, Min. Release = 13.5 MW.
Max. Storage = 157.7 GWh.

Time	Hydro Generation - 3 Load Segments			Storages	Inflows
1	99.9	14.3	14.5	76.0	19.8
2	111.5	53.2	14.5	76.9	43.2
3	111.5	111.3	14.6	80.6	50.7
4	111.5	111.4	14.6	70.9	70.7
5	111.5	111.4	14.5	74.5	74.4
6	111.5	110.0	14.5	80.6	66.0
7	111.4	95.0	14.5	81.6	46.7
8	108.1	27.9	14.5	74.3	39.0
9	90.3	14.5	14.5	83.4	28.6
10	15.1	14.7	14.5	90.5	21.1
11	46.7	14.6	14.6	94.8	15.2
12	72.1	14.0	14.6	94.2	16.2
13	107.0	65.0	14.8	93.6	15.5
14				75.1	

Reservoir: Cobb. Max. Release = 27.0 MW, Min. Release = 0.
Max. Storage = 29.4 GWh.

Time	Hydro Generation - 3 Load Segments			Storages	Inflows
1	19.5	18.8	16.8	29.4	12.3
2	20.5	22.5	19.7	25.6	21.3
3	25.3	23.0	23.8	25.8	26.3
4	26.9	25.8	21.7	27.7	19.1
5	26.2	22.7	25.0	24.4	24.6
6	26.2	22.0	24.7	24.8	20.9
7	25.5	18.7	22.4	23.0	24.4
8	17.1	20.0	17.8	25.4	19.4
9	5.7	22.3	9.9	25.7	18.7
10	12.9	18.5	6.9	27.8	14.7
11	21.7	12.6	5.3	29.2	9.7
12	3.2	9.4	5.9	29.3	7.5
13	11.7	18.0	9.9	29.3	9.9
14				26.6	

TABLE 6.5(c)

Reservoir: Coleridge. Max. Release = 31.8 MW, Min. Release = 0.
Max. Storage = 38.7 GWh.

Time	Hydro Generation - 3 Load Segments			Storages	Inflows
1	29.9	30.1	30.6	38.7	23.5
2	29.3	30.1	30.6	34.1	27.4
3	14.9	29.0	30.3	32.1	27.3
4	31.8	29.0	30.3	30.9	23.8
5	30.9	27.4	30.6	26.9	23.3
6	26.6	29.2	30.5	23.1	24.3
7	23.4	29.8	30.5	19.4	32.8
8	28.6	30.1	30.5	21.5	39.8
9	29.3	30.1	30.3	27.8	41.2
10	29.9	30.1	30.2	35.3	35.2
11	28.7	29.0	28.5	38.5	28.8
12	26.3	28.5	28.4	38.5	25.9
13	25.4	26.9	28.3	36.8	23.7
14				34.3	

Reservoir: Tekapo. Max. Release = 145.0 M, Min. Release = 0.

Time	Hydro Generation - 3 Load Segments			Storage	Maximum Storage	Inflows
1	119.7	92.3	130.6	290.0	294.9	110.4
2	24.6	114.8	136.8	289.0	307.2	90.4
3	136.2	38.1	132.3	268.4	319.4	77.4
4	144.9	139.3	129.5	261.5	331.7	57.2
5	140.1	77.8	135.5	209.3	331.7	56.5
6	141.7	37.9	134.3	174.5	307.2	55.0
7	3.0	25.2	127.8	151.8	294.2	78.3
8	6.2	103.8	130.8	155.4	283.0	113.5
9	38.6	104.1	118.2	156.4	283.0	149.8
10	4.3	118.9	115.6	184.8	283.0	167.6
11	110.9	119.8	120.0	222.2	283.0	164.3
12	19.0	115.1	131.0	252.3	283.0	156.0
13	1.7	111.7	129.7	277.8	283.0	131.6
14				289.0	294.9	

TABLE 6.5(d). Note: Storages in GWh. All other quantities MW averages.

Reservoir: Manapouri. Max. Release = 600 MW, Min. Release = 100 MW.
Max. Storage = 380 GWh.

Time	Hydro Generation - 3 Load Segments			Storages	Inflows
1	269.1	551.6	505.7	300.0	524.7
2	478.5	548.5	527.9	305.6	548.8
3	577.4	560.6	562.1	314.6	505.8
4	599.8	582.8	580.4	276.7	475.1
5	586.6	576.0	577.8	204.6	436.3
6	568.6	552.4	575.2	109.8	450.9
7	588.7	530.5	574.0	33.8	614.7
8	510.6	551.8	558.0	74.6	670.3
9	268.4	566.7	527.2	153.6	678.8
10	474.6	556.1	511.1	251.2	606.1
11	571.9	530.0	492.7	301.8	541.9
12	276.7	529.4	510.5	320.4	500.5
13	396.1	532.0	503.0	315.1	482.1
14				295.2	

Reservoir: Hawea. Max. Release = 78.4 MW, Min. Release = 2.3 MW.
Max. Storage = 234 GWh.

Time	Lake Releases - 3 Load Segments			Tributary	Storage	Inflows
1	77.5	13.9	11.2	134.0	230.0	18.6
2	75.7	21.3	9.6	133.1	232.0	19.1
3	60.8	69.7	23.3	120.6	232.5	17.4
4	78.3	72.6	18.2	101.4	212.5	15.0
5	75.0	71.5	34.4	92.3	191.0	15.0
6	70.5	72.8	34.4	96.0	164.8	15.8
7	60.8	10.9	14.3	130.1	138.9	23.4
8	70.4	6.3	5.4	164.5	144.6	30.0
9	76.1	4.5	4.1	188.6	158.8	32.6
10	72.7	5.7	4.1	181.8	175.5	28.3
11	72.4	5.6	4.2	163.4	189.1	23.4
12	75.3	7.8	6.2	146.6	199.3	20.2
13	75.6	10.0	11.1	139.6	206.0	18.7
14					209.4	

TABLE 6.5(e). Note: Storages in GWh. All other quantities MW averages.

Reservoir: Pukaki. Min. Release = 0. Max. Storage = 1091 GWh.

Time	Lake Releases - 3 Load Segments			Maximum Release	Tributary	Storage	Inflow
1	731.9	449.8	98.6	749.0	159.8	1000.0	242.9
2	734.9	542.8	93.3	749.0	142.2	1053.1	163.1
3	710.8	727.9	287.3	749.0	120.8	1031.7	126
4	748.7	721.1	455.5	749.0	93.7	838.5	97.1
5	723.8	708.4	338.9	749.0	83.7	610.4	93.3
6	699.2	720.3	305.3	749.0	84.5	400.9	90.4
7	708.7	698.8	190.5	749.0	115.6	181.6	127.8
8	587.9	453.0	118.7	749.0	174.2	18.2	184.0
9	659.1	316.8	106.6	712.3	232.9	29.1	254.3
10	422.6	197.7	113.0	692.0	259.0	129.0	323.8
11	166.9	176.4	92.5	709.8	236.1	322.6	384.5
12	687.8	253.4	111.1	724.9	216.6	584.5	387.3
13	647.2	314.5	137.8	749.0	183.2	800.0	323.2
14						943.1	

D.C. Link Power Flow at Sending End (South Island)

Max. Power = 601.4 MW.

Time	Peak	Intermediate	Off Peak
1	600.1	564.3	370.1
2	601.4	598.1	335.1
3	601.9	601.6	444.1
4	584.0	602.2	524.6
5	603.7	602.1	477.1
6	602.7	601.0	462.7
7	602.2	599.4	419.2
8	600.5	601.1	447.6
9	594.5	601.0	490.9
10	583.3	546.6	526.9
11	540.6	500.6	486.6
12	601.5	492.9	472.7
13	586.4	489.6	440.4

TABLE 6.5(f). All quantities MW averages.

Thermal Station Outputs (Peak, Intermediate and Off Peak).

	New Plymouth Max. Generation = 314.7			Stratford Max. Generation = 186.6		
1	314.7	273.3	62.8	120.0	0	0
2	314.7	314.7	56.8	186.6	42.8	0
3	314.7	314.7	100.5	186.6	186.6	0
4	314.7	314.7	93.9	186.6	186.6	0
5	314.7	314.7	62.4	186.6	186.6	0
6	314.7	314.7	65.2	186.6	186.6	0
7	314.7	314.7	58.5	186.6	106.2	0
8	314.7	314.7	59.9	186.6	81.4	0
9	314.7	313.8	80.2	89.6	0	0
10	314.7	170.4	104.2	27.2	0	0
11	242.8	168.6	158.1	0	0	0
12	314.7	298.9	133.6	4.5	0	0
13	314.7	305.7	122.3	82.7	0	0

	MEREMERE Max. Gen. = 78.3		MARSDEN Max. Gen. = 177.1		WHIRINAKI Max. Gen. = 38.9	OTAHUHU Max. Gen. = 59.3	Shortfall
	Peak	Intermediate	Peak	Intermediate	Peak	Peak	Peak
1	0	0	0	0	0	0	0
2	78.3	0	8.3	0	0	0	0
3	78.3	74.4	177.1	0	38.9	17.3	0
4	78.3	78.3	177.1	109.7	38.9	59.3	68.7
5	78.3	78.3	177.1	45.2	38.9	59.3	40.7
6	78.3	20.2	177.1	0	38.9	4.4	0
7	78.3	0	85.6	0	0	0	0
8	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0

No generation from Meremere, Marsden, Whirinaki, Otahuhu for load segments not shown.

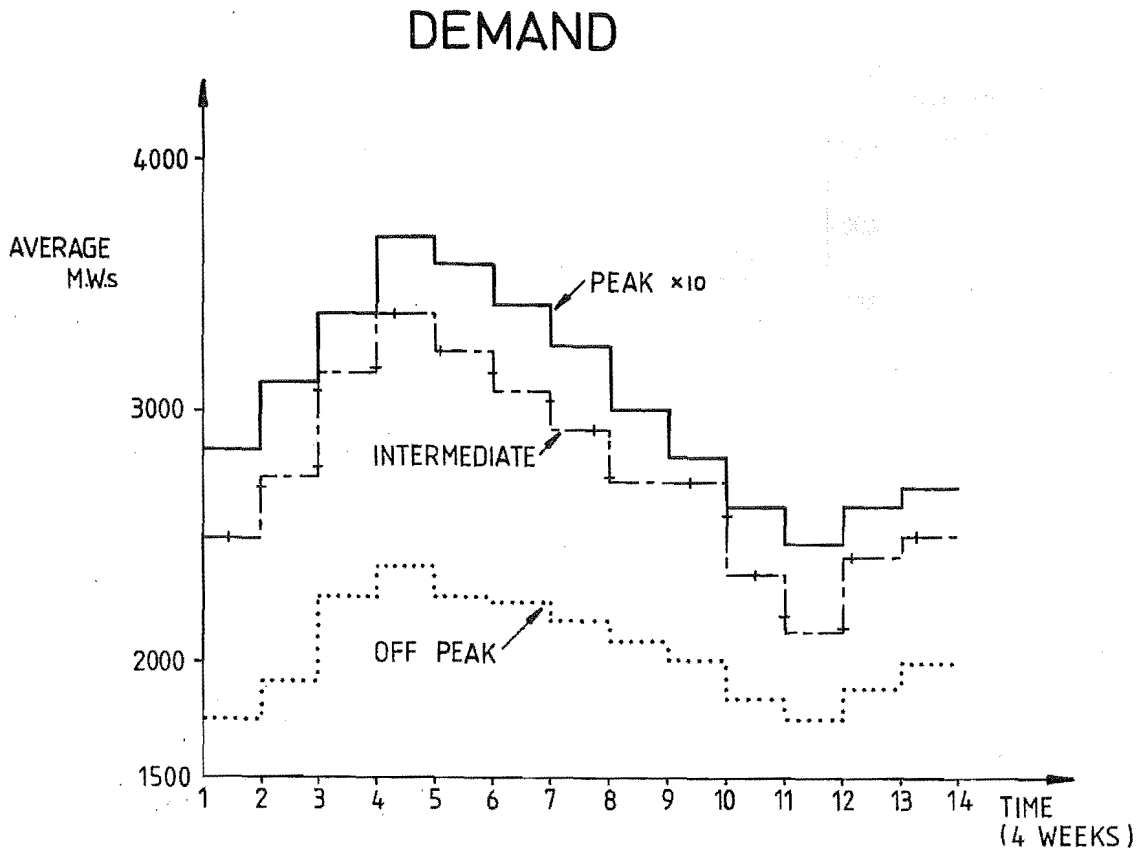
TABLE 6.5(g)

Incremental Costs ($/MWh)

Time	1	2	3	4	5	6	7
Peak	16.9	73.2	120.0	200.0	200.0	117.3	83.4
Intermediate	13.5	16.9	42.93	83.4	83.4	22.0	16.9
Off Peak	13.5	13.5	13.5	13.5	13.5	13.5	13.5

Time	8	9	10	11	12	13
Peak	19.45	16.9	16.9	13.5	15.81	16.9
Intermediate	16.9	15.08	13.5	13.5	13.5	13.98
Off Peak	13.5	13.5	13.5	13.5	13.5	13.5

Station	Incremental Cost ($/MWh)
New Plymouth	13.50
Stratford	16.90
Meremere	22.00
Marsden	83.40
Whirinaki	111.8
Otahuhu	120.0
"Shortfall"	200.0



KEY FOR RESERVOIR RELEASE and GENERATION GRAPHS :

PEAK ——————

INTERMEDIATE —+——+——+—

OFF PEAK .....

FIGURE 6.1(a) Results for year ended 31 March 1982 Optimal Simulation Also tabulated in Table 5.5.

TAUPO

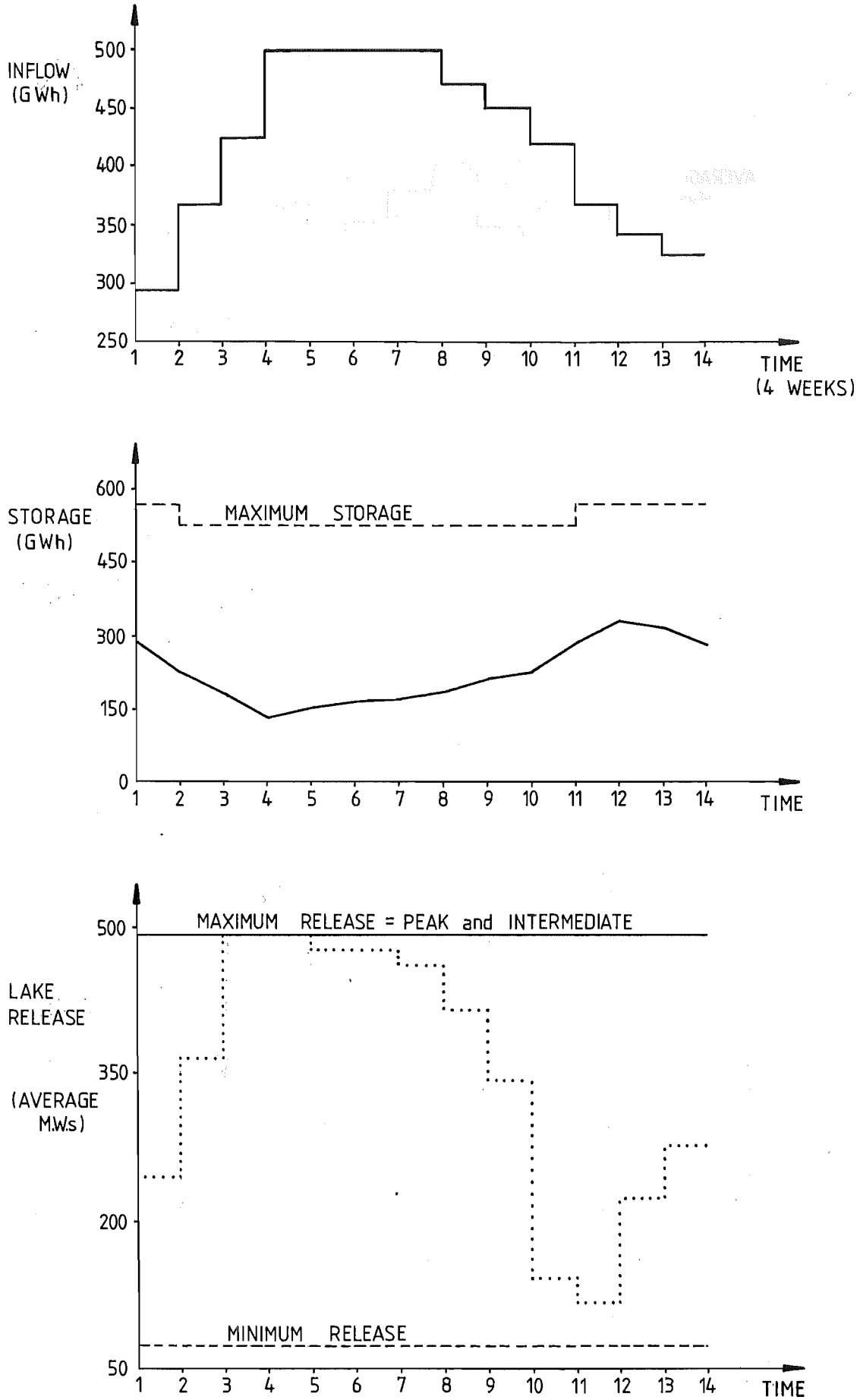


FIGURE 6.1(b) 1982 Optimisation (Continued)

WAIKAREMOANA

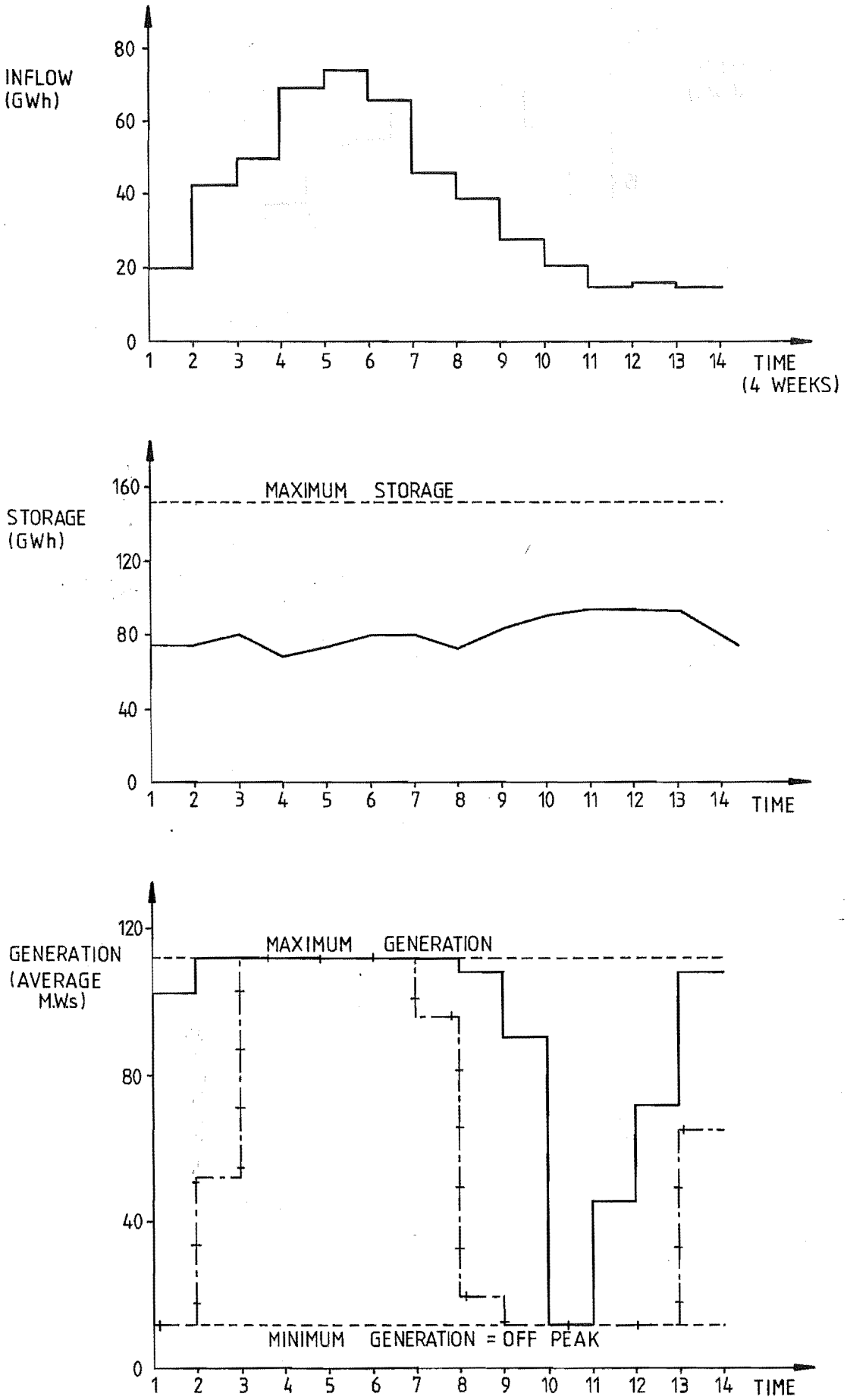


FIGURE 6.1(c) 1982 Optimisation (Continued)

COBB

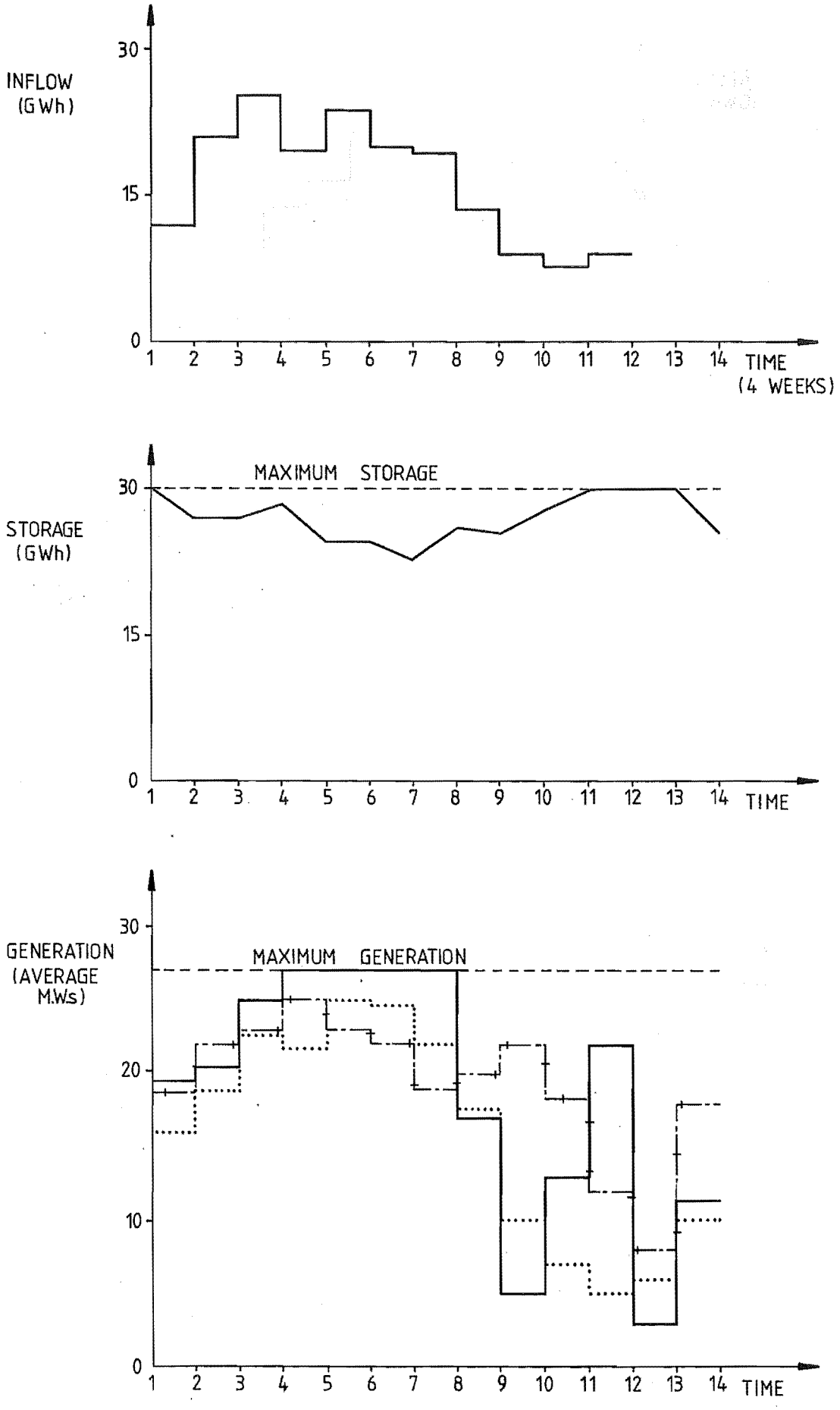


FIGURE 6.1(d) 1982 Optimisation (Continued)

COLERIDGE

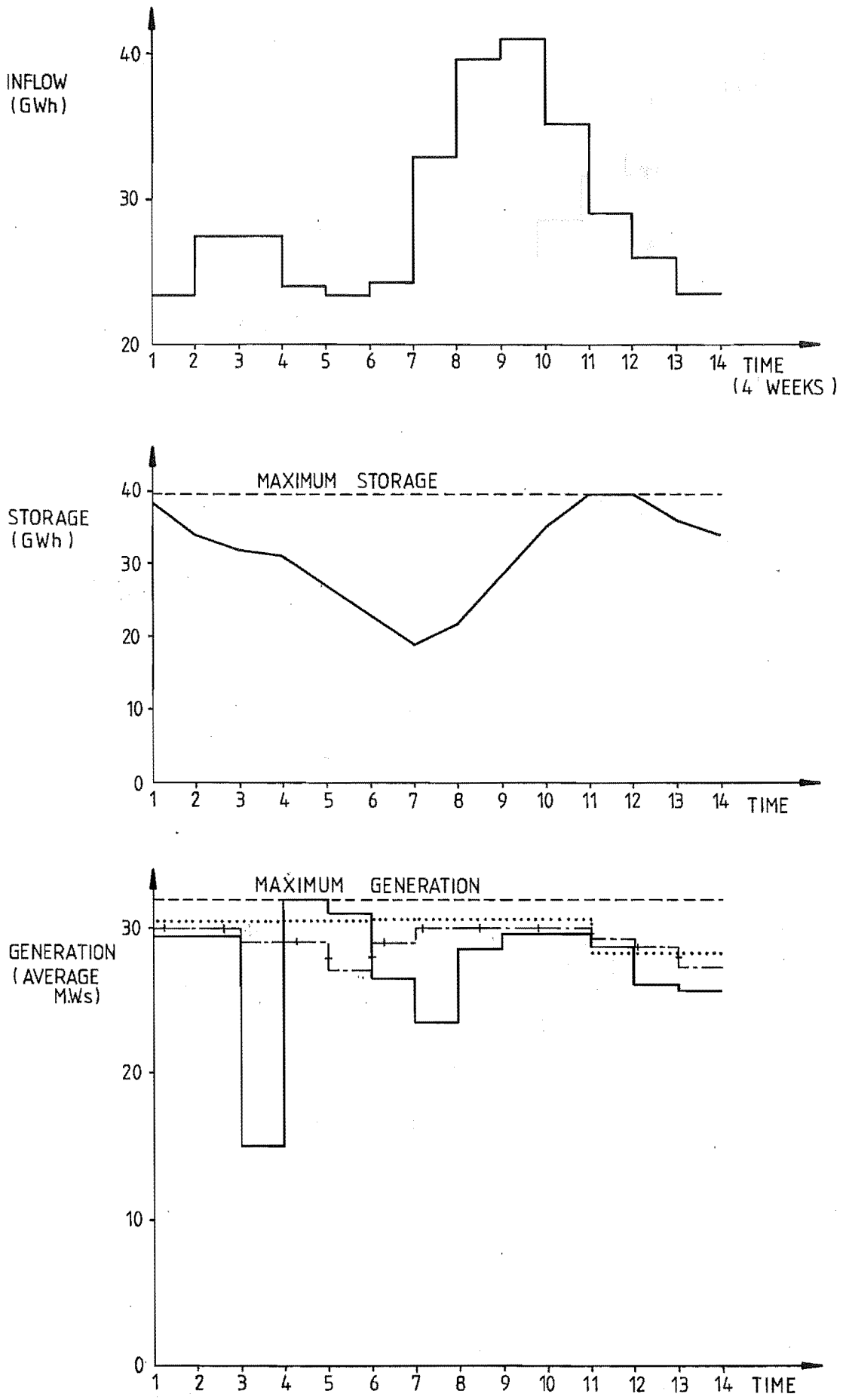


FIGURE 6.1(e) 1982 Optimisation (Continued)

TEKAPO

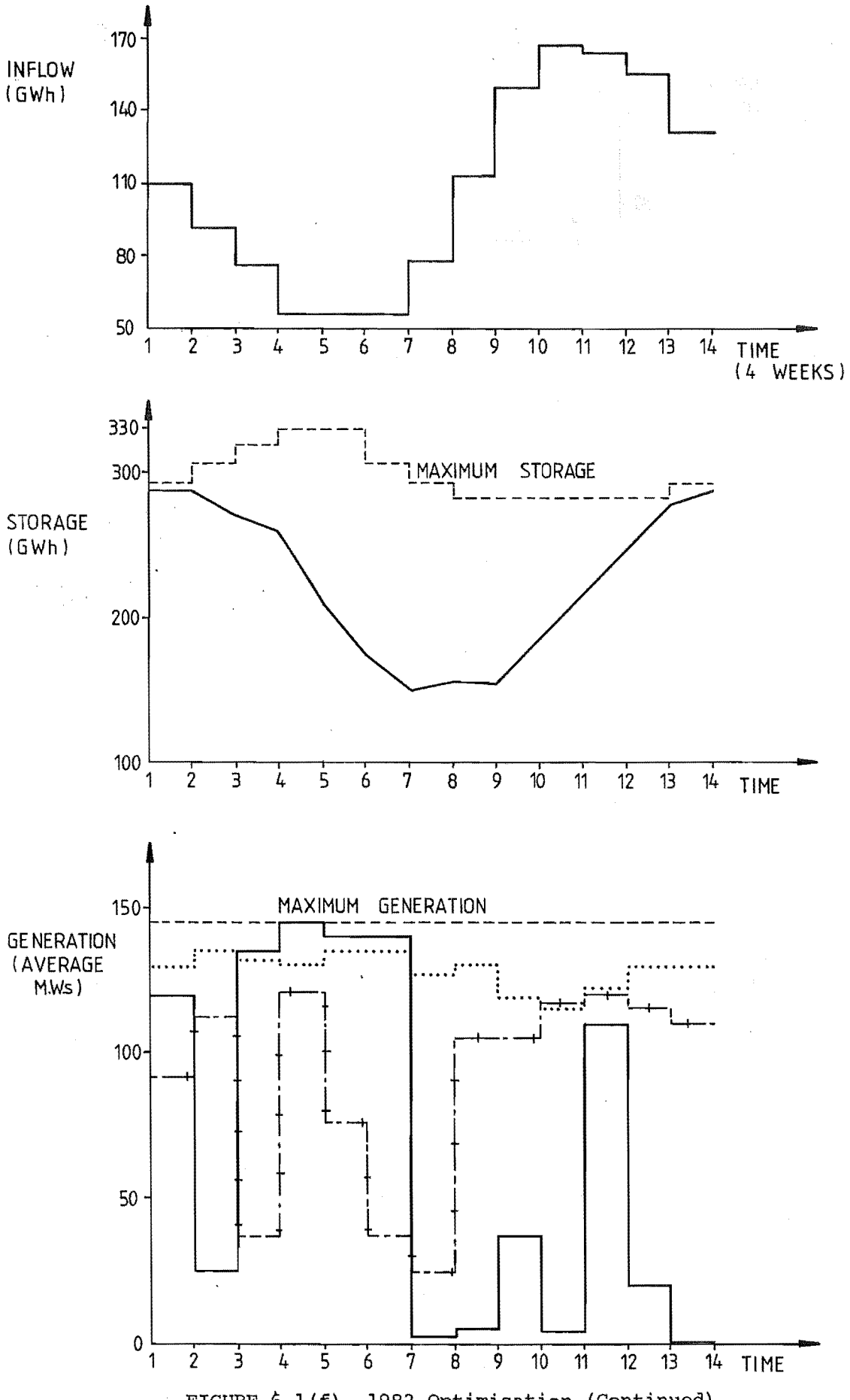


FIGURE 6.1(f) 1982 Optimisation (Continued)

MANAPOURI

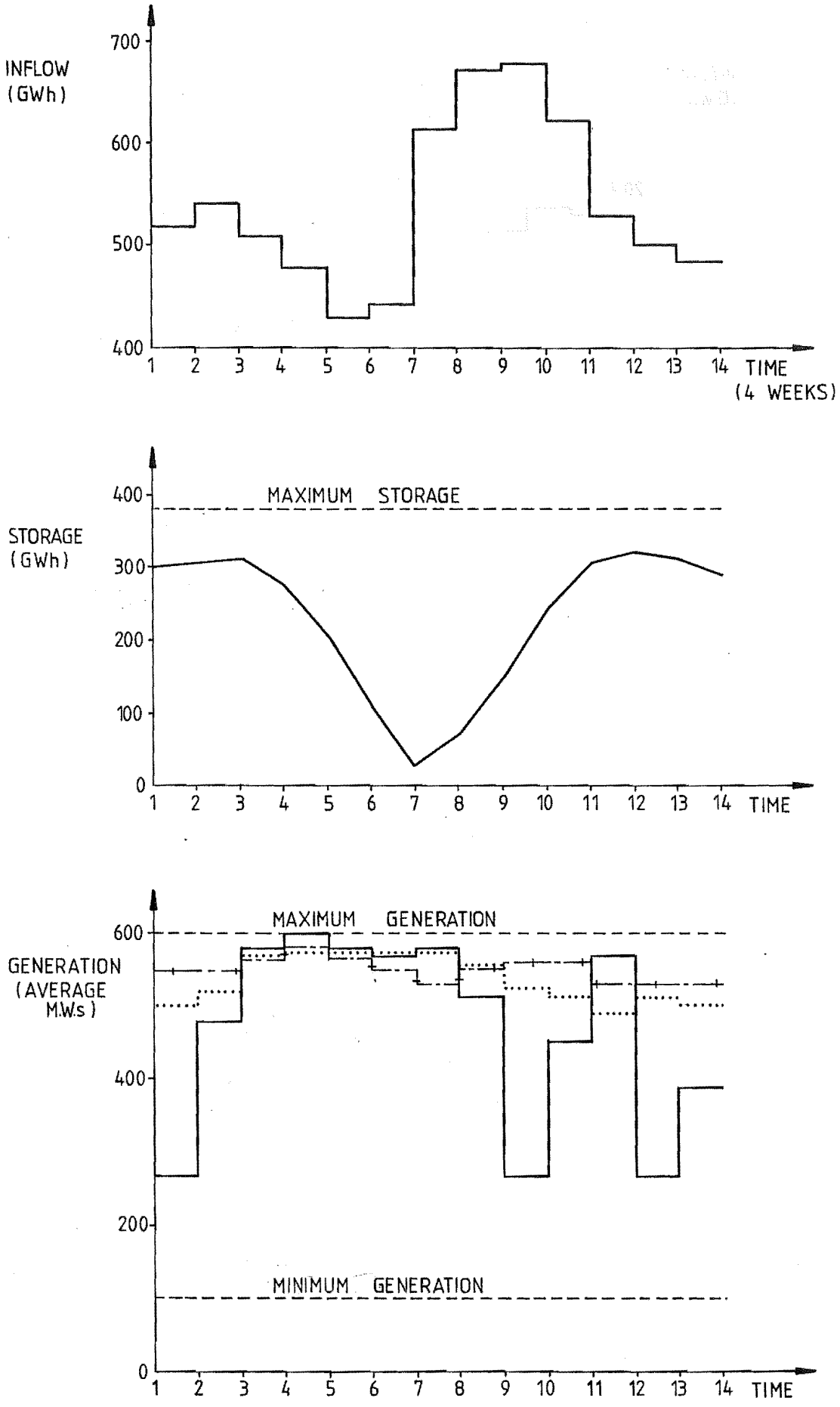


FIGURE 6.1(g) 1982 Optimisation (Continued)

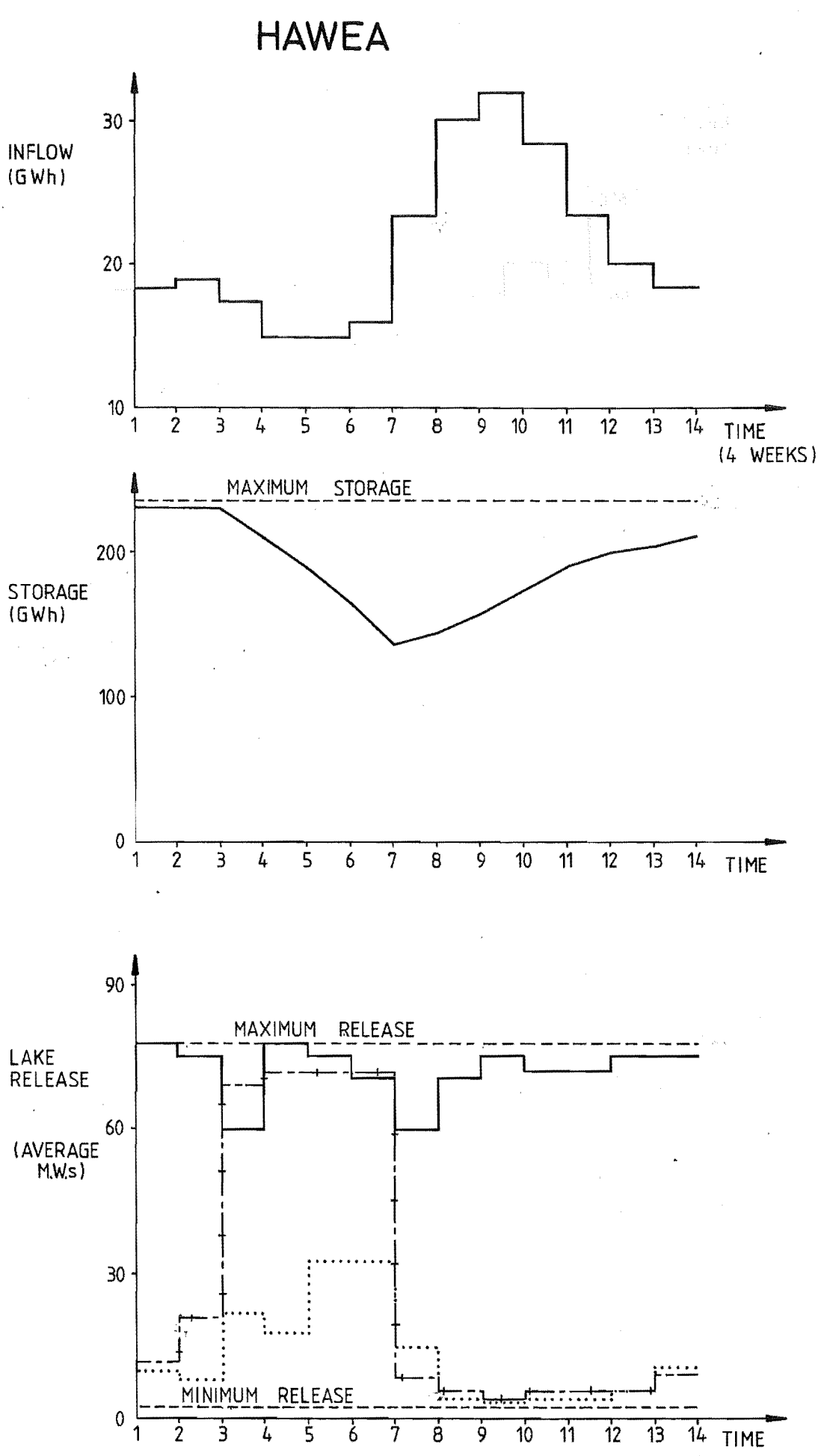


FIGURE 6.1(h) 1982 Optimisation (Continued)

PUKAKI

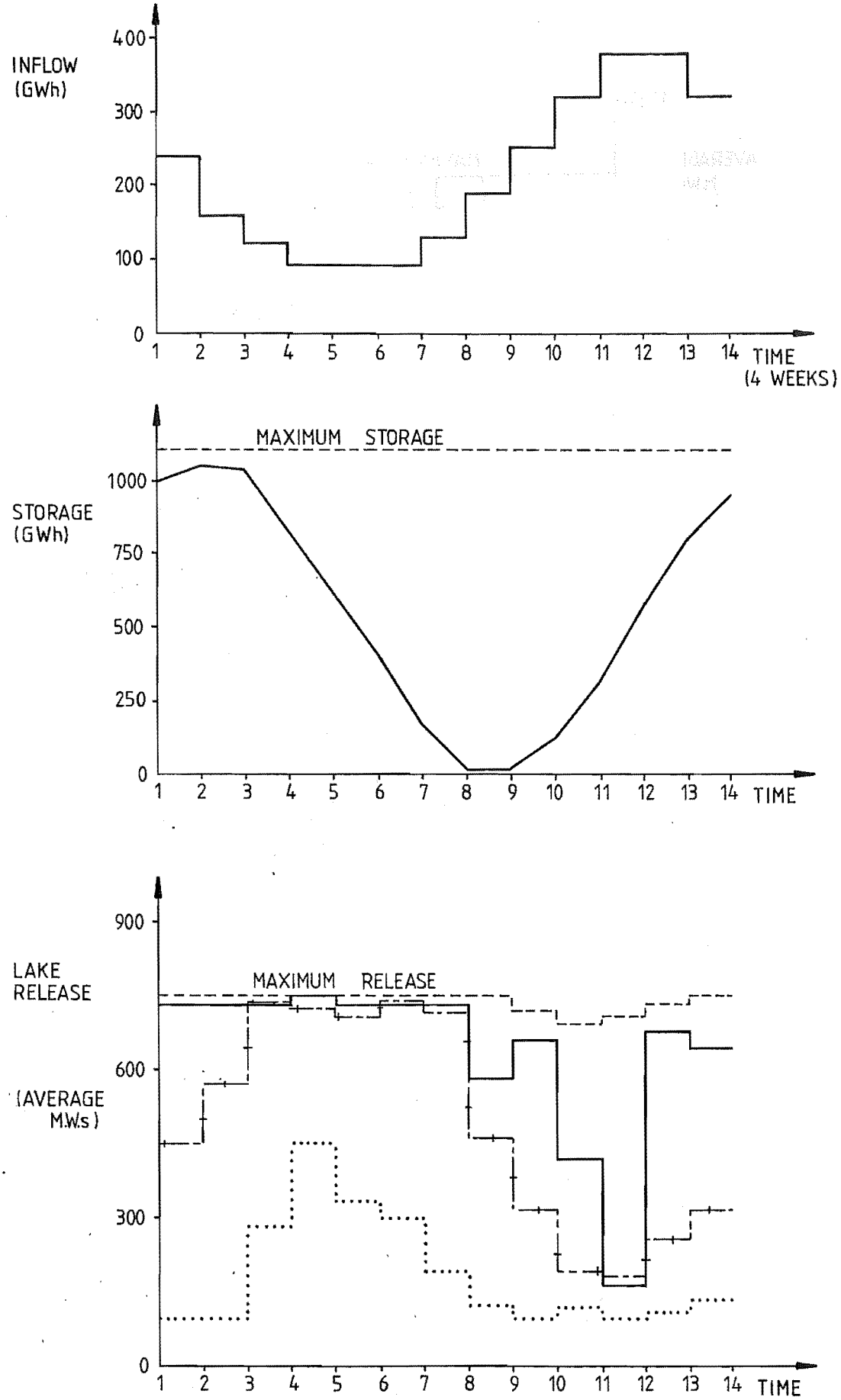


FIGURE 6.1(i) 1982 Optimisation (Continued)

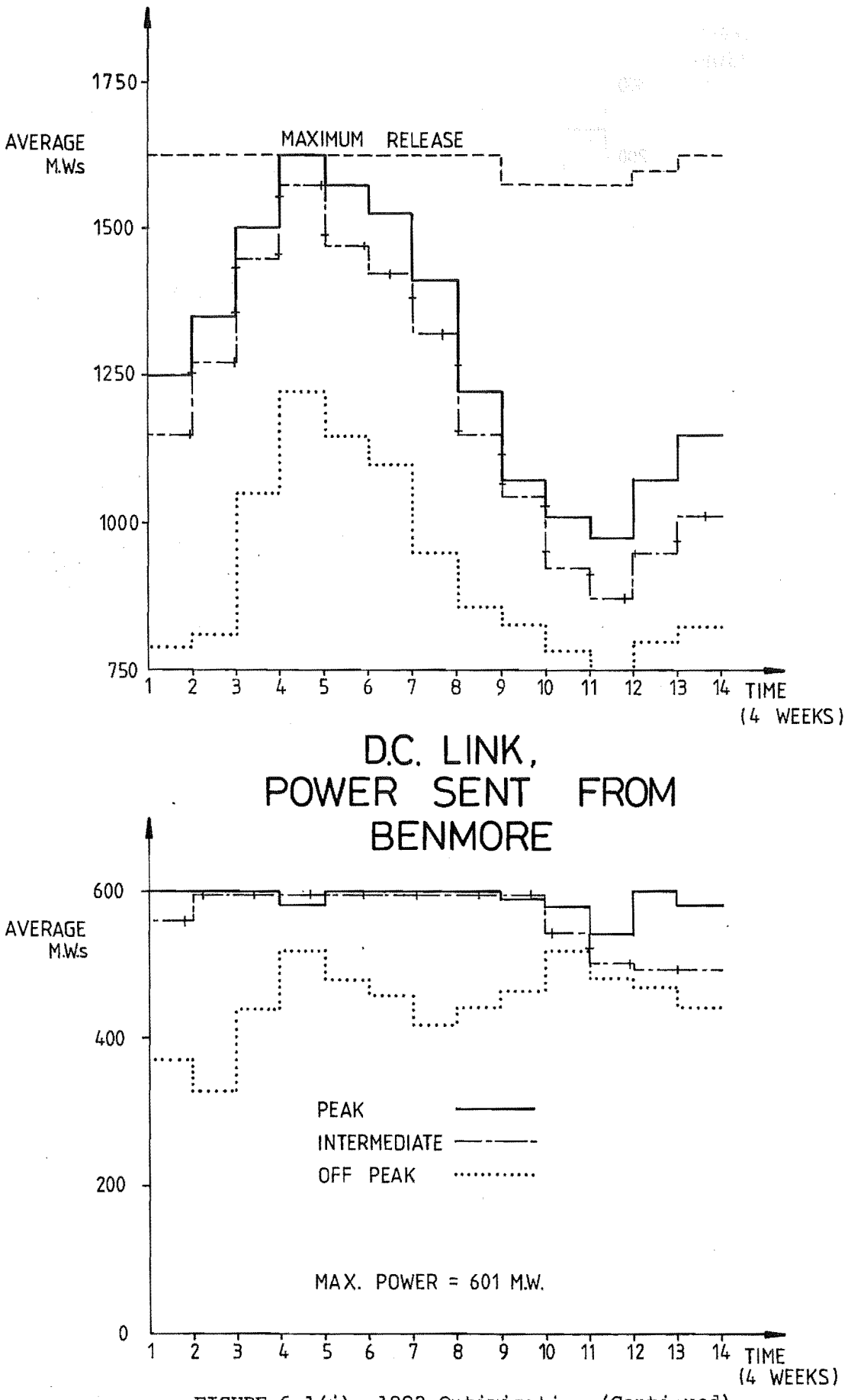


FIGURE 6.1(j) 1982 Optimisation (Continued)

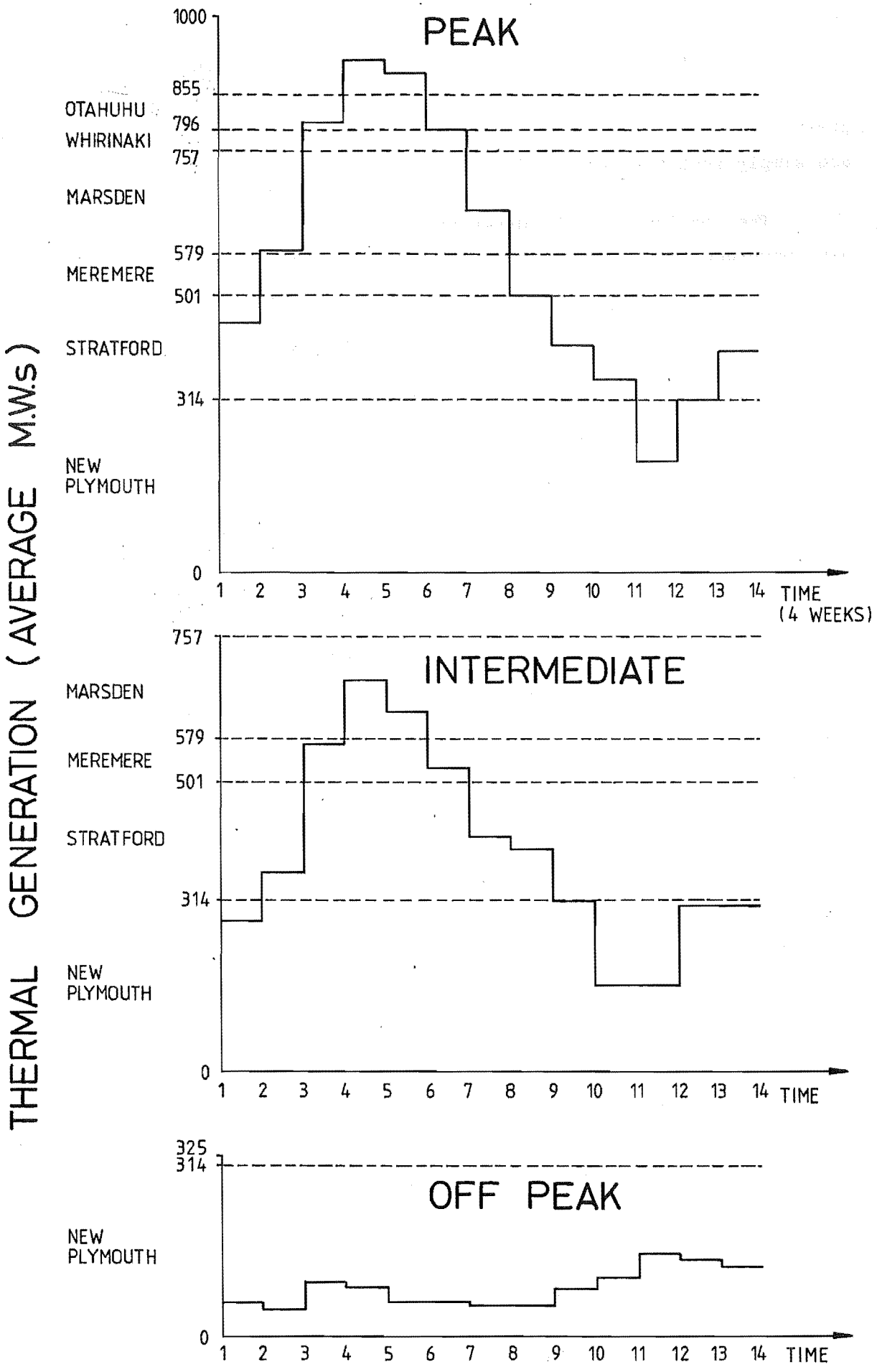


FIGURE 6.1(k) 1982 Optimisation (Continued)

maximum. This corresponds to July, when South Island demand is too high to allow a greater transfer north. All other transfer levels for costs greater than 13.5 exceed 97% of capacity. For the remaining times transfers are simply such that no inflows are wasted.

The conclusion to be drawn from this section is that the algorithm has converged to a solution which is valid except for one release from the Waikaremoana reservoir. This figure, 34% of maximum as against 1.37% expected, increases the cost at the solution by about 1 part in 10^4 ($3,480).

6.3.4 Comments on 1982 Operating Pattern

Operating patterns are shown best by the graphs of Figure 6.1. The demand graph (Figure 6.1(a)) shows how all three load segments peak for time = 4, i.e. late June and July. Demand for time = 5 is only slightly lower for the peak load segment. Taupo inflows (Figure 6.1(b)) conveniently peak over the range t=4 to t=7, with storage reaching its minimum level at t=4. So Taupo inflows are in phase with demands. Waikaremoana inflows (Figure 6.1(c) are also in phase, peaking at t=5.

Storages in the major South Island lakes - Tekapo, Manapouri, Hawea and Pukaki - all follow a similar pattern. Inflows peak at about t=10, 6 time intervals after maximum demand. Lake levels all reach a minimum either as inflows begin to increase again after the winter, with spring rains and snow melt, or one time interval later. In contrast to the winter rain fed North Island lakes, the South Island inflows are out of phase with demand, giving a strong cyclic pattern to storage levels. Figure 6.1(j) shows how the delay effect of storage reservoirs allows total South Island generation to peak at t=4 and follow the same pattern as the demand.

Figure 6.1(k) of thermal generation indicates how the incremental costs vary, following demand patterns. For the off-peak periods which all have the same incremental fuel costs, the quadratically increasing transmission losses will tend to smooth the pattern from one time interval to the next. If there were no losses, then costs would not be affected if generation were to be increased in one time interval, and decreased the same amount in another, provided fuel costs were unchanged.

TABLE 6.6 Transmission Losses as Percentage of Generation

Load Segment	1981	1982
Peak	4.70	4.49
Intermediate	5.15	4.78
Off Peak	5.53	5.00

Table 6.6 shows transmission losses for the 1981 and 1982 simulations. Lower percentage losses for the peak load segments could be due to a greater proportion of energy being supplied by stations close to the main North Island load centres, i.e. all the thermal stations are in the north, so at peaks a smaller proportion of power comes over the D.C. link.

6.4 EFFECTS OF SOME MODIFICATIONS TO THE SYSTEM

The deterministic scheduling method might be useful as a means of simulating an optimal operating strategy with various modifications to the power system. This could assist with the best use of capital for system expansion, or help show up system weaknesses. It might be more readily accepted in such applications than as an aid to system operation. With system expansion problems, uncertainty is greater, and the scope for experienced judgement less.

Three main aspects have been tested - various water inflow levels, inter island D.C. link capacity expansion, and the effect of a large new load on the year ended 31 March 1981 case. This last test might be typical of those helpful in energy pricing negotiations with new customers.

The addition of a 1500 GWh per year load at the Invercargill bus simulates approximately the effect of the new third potline at the Tiwai Point aluminium smelter. With the five year average flows, the average cost of the additional fuel needed to supply this load is 11.29 $/MWh (1.13 cents/KWh). However the cheapest thermal station has an incremental cost of 13.5 $/MWh, and no water wastage through spill occurs even without this extra load.

So the average cost of supplying this load is lower than the cheapest thermal station's incremental cost. This is due to the reduction in transmission losses from an average of 5.05% to 3.25%. The additional South Island load is compensated for by North Island thermal generation, reducing D.C. link losses, which are set at 10% of energy sent. A shortfall of 1.45 GWh occurs in one peak load segment. Both Otahuhu and Whirinaki are in use for two peak load segments unlike the case without the extra load, where Whirinaki is the most costly station used.

TABLE 6.7 Effect of Various Inflow Levels. Table of Annual Load Factors (%age Utilisation of Available Plant) for Thermal Power Stations.

	620 MW D.C. Link							800 MW D.C. Link	
Flows: x 5 Year Av.	1.2	1.1	1.05	1.0	0.95	0.9	0.8	1.0	1.1
New Plymouth	31.3	34.4	35.8	40.3	63.6	85.5	100.0	45.3	23.0
Stratford	12.6	14.3	14.9	15.2	17.9	18.7	90.1	7.6	5.7
Meremere	2.5	3.5	4.1	4.9	6.5	6.6	11.7	1.4	1.1
Marsden	0.6	0.9	1.0	1.1	1.2	1.2	1.4	0.5	0.4
Whiriraki	0	0	0	0	0.3	0.7	0.7	0	0
Otahuhu	0	0	0	0	0	0.3	0.4	0	0
Shortfall	0	0	0	0	0	0	0	0	0
Total Cost ($ Millions)	17.84	20.20	21.16	23.45	34.56	44.21	70.15	22.20	11.95

Table 6.7 summarises the other two series of tests. A 10% reduction in inflows causes an 88% increase in fuel cost, whereas a 10% increase in flows saves only 14% on fuel costs. Obviously it is more important to minimise the effect of low flows than to take advantage of high flows. 10% of total flows corresponds to about 1.45 standard deviations, assuming a normal distribution, so it is a fairly severe case. Even with 20% low inflows, no shortfall in supply occurs - just a large increase in generation at New Plymouth.

In the 10% high flows case, about 4.8% of inflows are spilled, almost all in the South Island. Under these flow conditions, increasing the D.C. link capacity to 800 MW approximately halves the fuel cost, a reduction of $11.5 million per year. Spilling drops to 1.9% of inflows, all in the South Island. For the five year average flows however, the saving in fuel cost due to the larger D.C. link capacity is relatively small - only 5.3% or $1.25 million. The additional capacity is only used for some peak load segments, having a small effect on total cost. This suggests that under the conditions modelled, the choice of a 620 MW link is a good one. The 800 MW link is helpful only in high inflow years.

6.5 COMPARISON OF VARIOUS CONSTRAINT ENFORCEMENT METHODS

6.5.1 Introduction

The ability of CGRADS to solve the hydro-thermal problem with three constraint enforcement strategies was mentioned in section 4.6. Here the merits of the three are evaluated, and then the use of different types of penalty function is investigated. Only penalties of the form $(\underline{X}_v - X_v(t))^2$ have been described so far. Here powers other than 2 are tested, to see if an improvement in efficiency is possible.

6.5.2 Control Constraint Methods

Table 6.8 summarises results for the three control constraint enforcement techniques for the three hill climbing methods used.

Insofar as process times are concerned, there is little between the three constraint strategies, using CGRADS. The factor of two saving in process time using XLS, RSC parameter set (b) does not seem of such great consequence compared with the extra complexity and programming time required to include the scaling factors. The slight improvement in accuracy may be worthwhile.

TABLE 6.8 Comparison of 3 Hill Climbing Algorithms With 3 Control Constraint Enforcement Strategies.
(Thermal power costs relative to $\$0.235 \times 10^8$)

Control Constraints Enforced By:	Scaled Transformations		Unscaled Transformations		Penalty Functions	
Hill Climbing Method	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost
CGRADS	12.4	1.00 (a)	15.4	1.002 (a)	10.8	0.992 (c)
CGRADS	18.3	0.997 (b)	36.8	1.000 (b)	-	-
Steepest Descent	21.8	0.998	129.5	1.004	23.8	0.995
FMCG	265	1.000	Failed		Failed	

(a) XLS = 0.2 RSC = 0.5

(b) XLS = 0.05 RSC = 0.5

(c) XLS = 0.1 RSC = 0.5

where XLS = linear search accuracy

RSC = restart parameter.

The penalty function approach is considerably easier to program than either of the other two. Control and state initial penalty weights ranging from 1×10^{-3} to 1×10^{-5} were tested, all giving satisfactory results. The cost figures in Table 6.9 are lower for this method as a 2% tolerance on constraints allows extra generation above actual limits. In contrast exact limits are imposed by the transformations.

Fears of difficulty in choosing initial penalty weights and of difficulties arising from interaction of control and state penalties were felt when first choosing constraint enforcement methods. Hence control constraints transformations were first experimented with. This meant that trial and error investigations to find suitable initial penalty weights only involved the one type of variable (state penalty weights), rather than the two variable search needed if control and state penalties were used. It

turned out that no difficulties were encountered with the control penalties version, especially as we knew what state penalty weightings were suitable by that stage.

6.5.3 Penalty Function Tests

Four different penalty powers were tested: 2.0, 1.6, 1.2 and 1.1. Control constraints were enforced by scaled transformations. Results for the 1.1 penalty power are not shown in Table 6.9 with the others, as only one or two satisfactory runs were made.

A smaller penalty power imposes lower penalties on large constraint violations. This might help overcome the ill-conditioning that can be caused by penalty functions, especially if the initial weighting factor is too large. So fewer penalty iterations should be needed if a sufficiently large initial weight can be used with a lower penalty power.

Consider the square penalty, (Table 6.9(a)), with x10 weighting increases each iteration. Small initial weights require extra computation, but a 1×10^{-3} initial weight increases fuel costs by 1.7% over the (satisfactory) base value of $\$.235 \times 10^{8}$. Increasing weights by 100 each iteration is not satisfactory - solutions are obviously not sufficiently accurate. A smaller rate of increase, x3.2, gives the most accurate results, but can require more computation than the x10 strategy for some initial weights.

As the penalty power decreases, so should the rate of increase of weights. The range of weighting increase factors tested is therefore lower for the other two penalty powers shown in Table 6.9:

(a) For a square penalty, if

$$F(x) = f(x) + W(\underline{X}-x)^2$$

then $$F'(x) = f'(x) + 2W(\underline{X}-x)$$

$$= 0 \quad \text{at the solution}$$

$$\Rightarrow \quad W = \frac{-f'(x)}{2(\underline{X}-x)}$$

So on the next penalty iteration, increasing W by a factor of 10 should reduce $\underline{X}-x$ by about the same factor.

TABLE 6.9 Tests of Various Penalty Functions for State Constraints.
(Scaled Control Transformations)
Fuel costs are relative to $.235 \times 10^8$

(a) Square Penalty Function, $W(\underline{X}_v - X_v(t))^2$

Increase in Penalty Weighting Each Iteration	100		10		3.2	
Initial Penalty Weighting	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost
10^{-8}	33.6	1.006	44.2	0.996	58.1	0.989
10^{-7}	68.6	1.045	42.7	0.998	72.3	0.989
10^{-6}	34.7	1.003	33.7	1.000	54.5	0.992
10^{-5}	27.7	1.052	23.5	0.997	40.7	0.991
10^{-4}	11.8	1.049	19.0	0.999	15.6	0.995
10^{-3}	22.6	1.021	10.9	1.017	14.8	1.010
10^{-2}	6.5	1.105	7.6	1.101	-	1.095

(b) 1.5 Power Penalty Function, $W(\underline{X}-X_v(t))^{1.5}$

Increase in Penalty Weighting Each Iteration	10		3.2		1.8	
Initial Penalty Weighting	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost
1×10^{-5}	51.2	1.009	58.7	0.999	-	-
1×10^{-4}	34.5	1.006	65.1	0.999	83.4	0.998
1×10^{-3}	24.9	1.005	35.3	1.001	62.7	0.993
3.2×10^{-3}	26.3	1.011	31.0	0.998	22.0	0.989
1×10^{-2}	16.5	1.005	31.3	0.997	30.4	0.987
3.2×10^{-2}	10.6	1.015	12.1	0.995	21.9	0.989
1×10^{-1}	7.1	1.033	10.2	0.994	14.1	0.992
3.2×10^{-1}	5.3	1.166	9.2	1.126	10.9	1.064
1×10^{0}	4.6	1.362	-	1.362	-	1.362
1×10^{1}	-	-	-	-	-	-

... ...

TABLE 6.9 (Contd)

(c) 1.2 Power Penalty Function, $W(\underline{X}_v - X_v(t))^{1.2}$

Increase in Penalty Weighting Each Iteration	3.2		1.8		1.3	
Initial Penalty Weighting	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost	CPU (min.)	Thermal Power Cost
1×10^{-2}	68.2	1.016	81.3	1.030	-	-
3.2×10^{-2}	50.8	1.012	78.1	1.001	-	-
1×10^{-1}	26.4	1.019	44.9	1.001	76.7	1.007
1.6×10^{-1}	27.4	1.009	38.1	1.010	-	-
3.2×10^{-1}	22.1	1.010	33.9	1.001	44.5	0.992
5.62×10^{-1}	15.6	1.004	14.9	0.994	61.1	0.990
1×10^{0}	19.0	1.015	26.5	1.003	37.0	0.996
1.6×10^{0}	10.1	1.009	15.1	1.009	25.9	0.999
3.2×10^{0}	12.1	1.081	14.9	1.041	26.2	1.013
5.62×10^{0}	-	-	8.9	1.066	9.4	1.064
1×10^{1}	-	-	9.0	1.089	8.7	1.102

(b) For a 1.5 power penalty,

$$F'(x) = f'(x) + 1.5W(\underline{X}-x)^{\frac{1}{2}}$$

$$\Rightarrow W = \frac{-f'(x)}{1.5(\underline{X}-x)^{\frac{1}{2}}} \quad \text{at the solution}$$

So to reduce $\underline{X}-x$ by 10, the value of W need be increased by only $\sqrt{10}$.

In Table 6.9(b) for the 1.5 penalty power, it is x10 increases in weights that give unsatisfactory results (as opposed to x100 for 2.0 power penalty). Again, the slowest rate of increase of weights, x1.8, gives the most accurate results. For a 1.2 penalty power, 1.3 increase in weights

is most reliable, but process times tend to be higher than for the other penalty powers.

From these tests it can be seen that no great advantage is obtained from the use of other than a square law penalty. The most striking conclusion is that choice of a suitable increase in penalty weighting each iteration is most important. In some cases shorter process time and more accurate solutions are sometimes possible simply by using a smaller rate of increase.

Smaller penalty powers give reasonable accuracy and process times for a more limited range of initial weights. The trial and error determination of satisfactory initial weights will therefore be more time consuming. The 1.5 power penalty, with 1.8 increases in weighting factors, gave generally more accurate solutions than the 2.0 power penalty, and still with reasonable process times. These reductions in cost are made by setting variables closer to their allowable limits, these limits being the same for all runs of course.

6.6 CONCLUSIONS

Optimisation runs of the New Zealand system model, approximating some aspects of actual operations for N.Z.E.'s 1981 and 1982 financial years have been examined. This provided a check on data accuracy, and the results presented show that there are no gross errors in the model. Fuel costs and reservoir costate values at the solution have been used to justify the release strategy. Only one inexplicable feature could be found, giving a fuel cost increase of 1 part in 10^4, (i.e. \$3,480). All other aspects of the solution seem reasonable.

The possible use of the algorithm as a system development tool has been demonstrated. A larger D.C. link capacity was shown to be of little benefit, with the demand data etc. used. The addition of the Tiwai Point aluminium smelter extension was simulated. Percentage transmission losses dropped, and the additional load had the surprisingly low incremental generation cost of 1.13 cents/KWh. (Note that fuel cost data is already outdated).

High water inflows were shown to produce smaller effects on costs than lower than mean inflows. This suggests that a stochastic method giving more conservative strategies (a lower risk of running out of water) might be of considerable benefit.

The control constraints enforced by transformations method involves more complex programming than the alternative penalty function method. Penalty functions do not enforce constraints exactly, hence a falsely lowered fuel cost figure is given. In practice generator capacities are a "hard" constraint, while reservoir levels are "softer" constraints, often limited by environmental requirements. This makes the release transformations, storage penalties combination seem appropriate.

No decisive argument for the use of penalty functions other than the square law form was found. A 1.5 power penalty, with a 1.8 increase factor applied to weights, produced the most accurate (i.e. cheapest) solutions, and with reasonable computer process times. Savings of the order of 0.5% are possible, and may justify this form. Choosing the rate at which scalar weighting factors increase was shown to be as important as the choice of initial weights.

CHAPTER 7

INTRODUCTION TO THE STOCHASTIC PROBLEM

7.1 INTRODUCTION

The solution method presented in the preceeding chapters made no allowance for reservoir inflow variation. These variations are quite large in New Zealand, even on the basis of total national inflows. Total monthly inflows have a standard deviation of about 30% of the mean while for total annual inflows the corresponding value is 7%.

Various aspects of the stochastic problem will be described in this chapter, leading up to the method of Chapter 8 which might provide a stochastic multi-reservoir scheduling solution. A multi-reservoir problem has not been solved here, but the demonstration on a simple problem of an algorithm, which has the potential to do so, is the objective.

Uncertainty makes system operators more cautious than if they had perfect knowledge of inflows. For example, storage lakes would not be deliberately allowed to empty completely at the end of winter in anticipation of large spring inflows. It is more reasonable to leave some margin in case spring flows are late arriving, or a dry spell occurs at the end of winter. The deterministic algorithm proposed earlier does not allow margins to guard against such events.

An exact solution to the stochastic problem can be obtained by stochastic dynamic programming (S.D.P.). The well known "curse of dimensionality" (Bellman, 1961) prevents its use for multi-reservoir problems, certainly with more than two reservoirs. In Chapter 2, some single reservoir S.D.P. solutions were mentioned, but such aggregation of several reservoirs into one is particularly unsuitable for the New Zealand system. Here, Taupo inflows are conveniently in phase with the winter peak in electricity demand. South Island inflows peak in spring, with lakes usually full at the end of summer. Hence management patterns are quite different and separate modelling is essential. A two reservoir S.D.P. might be of some use, and might be computationally feasible, with a coarse discretisation of states.

7.2 SOME POSSIBLE SOLUTION METHODS

The possibility of an exact solution by stochastic D.P. for a multi-reservoir system has been dismissed, but a one reservoir model solved in this way is described in Section 7.6. A sub-optimal solution method which is feasible for a multi reservoir problem is developed. It is tested on a single reservoir model and the solution compared with the exact S.D.P. solution of the same model. If the two methods give similar solutions, then the sub-optimal method should also give useful solutions for the real multi-reservoir system.

A deterministic algorithm could be used as an open loop feedback controller (O.L.F.C). The deterministic multi-reservoir algorithm presented is an open loop controller as its output is a set of releases that do not depend on observations of actual reservoir storages. There is no need to make such observations with a deterministic system. To apply the algorithm to a stochastic system, observations of storages could be made at the end of each week. These updated values would then be used to produce a new solution for the full twelve months. Only the first week's release decisions will be implemented, of course. This observation and repeat solution process provides the feedback action of the open loop feedback controller. The Certainty Equivalence Principle (Patchell and Jacobs, 1971) states that a linear quadratic gaussian (LQG) system has the same optimal control law as a deterministic system, using expected values of the random variables. The O.L.F.C. will give an exact solution to such a problem.

The hydro-thermal problem is not linear, or quadratic. Storage and release constraints in particular cause problems. For comparative purposes a one reservoir D.P. solution has been obtained using the same data as the S.D.P. (data from section 7.5) but using average inflows only. This will show how far from an optimal strategy is that given by an O.L.F.C.

Chance constraints can be incorporated in a non-linear programming method. This would steer reservoir storage trajectories away from their limits by modelling the probability distribution of storages and penalising any violations of constrain which have a probability greater than some prescribed amount. This will give a mor cautious strategy, but does not give releases as a function of storage.

Linear programming can take account of stochastic aspects either by two stag programming, or the chance constraints method (Vajda, 1972). In both cases, the linear programming problem can become very large as probability distributions are represented as a set of points, each adding an equation to the set to be solved. The two stage approach penalises violations of constraints. This is the "second stage" of the problem.

Dillon *et al* (1980) demonstrates both chance constrained and recourse action linear programming methods. A cost is applied to spills to ensure that they are minimised. This is unrealistic, as spilling is a necessary, unavoidable action when sufficiently large inflows occur. Penalising it may invalidate the solution. We wish to treat spill as a permissible action, simply clipping off the storage variables at their maximum values, in the same way as for the deterministic model.

This sets the modelling technique here apart from the usual approaches used in water resource optimisations where upper and lower state constraints can be treated in the same way by chance constraints. Also, the ability to model non-linear transmission losses is necessary. These requirements prevent the use of linear programming. Some means of determining the optimal release as a function of storage is also needed, in a non-linear model.

The model given in section 7.7 uses a linear feedback rule with chance constraints, and is solved by non-linear programming. Further development of this model is given in Chapter 8 - the "gaussian plus impulse" model. Some aspects of reservoir inflows will be discussed next, followed by the D.P. solutions needed for comparison with the alternative models proposed.

7.3 INFLOW MODELLING AND SERIAL CORRELATIONS

Water inflows are not an entirely random phenomenon. Patterns are evident, for example in monthly inflows, with a similar pattern being followed each year. Superimposed on these predictable patterns is a random element. The statistics of flows in successive time intervals are not the same, so their statistics are non-stationary.

A first order Markov model that can be used to represent monthly total inflows is, from Linsley *et al* (1958):

$$Q_{ij} = \bar{Q}_j + \rho_j \frac{\sigma_j}{\sigma_{j-1}} (Q_{i-1,j-1} - \bar{Q}_{j-1}) + t_i \sigma_j \sqrt{1-\rho_j^2} \qquad (7.1)$$

where j defines the month of the year, 1 to 12 each year,

i indexes months from the start of this particular sequence, 1 to n serially,

Q_{ij} = total monthly inflow,

$\bar{Q}_j$ = average inflow for month j,

ρ_j = lag 1 serial correlation coefficient,

t = random variate, zero mean, unit variance,

σ_j = standard deviation.

t_i is either normal or log-normal, the log-normal distribution emphasising low flow volumes. The deterministic components of this model include each month's mean, variance, and serial correlation coefficient. These may, however, have a long term oscillation or cyclic effect. Trends may exist. It is a variation in deterministic parameters due to changes in the catchment conditions, perhaps caused by land development modifying run-off rates.

Determining the parameters for the model is even more difficult than might appear. There are the usual reliability problems with measuring equipment over many years, as was mentioned with the inflow data for the eight reservoir deterministic model of the New Zealand system. In addition, the serial correlation of flows results in a degree of redundancy in information yielded by each hydrological event, i.e. the observations are non-random. So statistical parameters obtained from a sequence of events are less reliable than is indicated by the sequence length. Before using a non-zero serial correlation coefficient, the significance of the calculated value must be checked. An approximate formula for this is given by Matalas (1967). For example, at the 95% significance level, a coefficient of 0.3, obtained from less than about 25 years data, is not significantly different from zero.

When calculating parameters from measured data, the calculated variance must be reduced to account for the coupling effect of serial correlations. First ρ_j is found, from (Yevjevich, 1972):

$$\rho_j = \frac{\sum_{t=1}^{n} x_t\, x_{t+1} - \left(\sum_{t=1}^{n} x_t\right)^2/n}{\sum_{t=1}^{n} x_t^{\,2} - \left(\sum_{t=1}^{n} x_t\right)^2/n} \qquad (7.2)$$

for n samples.

Then the correction factor is calculated:

$$\psi_j = \frac{n-1/n}{1 - \frac{1-\rho_j}{n(1-\rho_j)^2} + \frac{2\rho_j(1-\rho_j^{\,n})}{n^2(1-\rho_j)^2}} \qquad (7.3)$$

and finally the variance is

$$\sigma_j^2 = \psi_j \frac{\sum_{t=1}^{n} x_t^2 - n\bar{x}_j^2}{n-1} \qquad (7.4)$$

where $\bar{x}_j$ is the mean for month j.

The inflow model, eqn (7.1), can be used to generate synthetic inflow sequences, extending the available data set, as was done by McKerchar (1971, 1975) to determine optimal reservoir operating strategies. He found the values of serial correlation coefficients for several flows into hydro electric lakes in New Zealand, as given in Table 7.1.

TABLE 7.1 Monthly Flows-Average Serial Correlation Coefficients (McKerchar, 1971)

	Data Set Length (Years)	Lag One Serial Correlation	Lag Two Serial Correlation
Lake Tekapo Inflow	40	0.379	-
Lake Pukaki Inflow	32	0.291	-
Lake Ohau Outflow	39	0.435	-
Ahuriri River, at Benmore	15	0.451	-
Lake Manapouri Outlet	35	0.587	0.25
Lake Te Anau Outlet	35	0.566	0.241

Even if ρ is small, its effect can be most significant, increasing the spreading of the total inflow as time passes, and decreasing the randomness of the process, e.g. low inflows will tend to be followed by more low inflows, increasing the possibility of a reservoir emptying completely. When formulating a decision rule for releases from hydro reservoirs, it could be useful to take advantage of the information given by past flows about future ones. This could be done using the rule:

$$u(t) = f(x(t), I(t-1)) \qquad (7.5)$$

where $x(t)$ is the present storage and $I(t-1)$ is the previous inflow. (or more specifically, from the knowledge of $I(t-1)-\bar{I}(t-1)$, where $\bar{I}(t-1)$ is the mean flow).

Maidment (1976) used stochastic D.P. to optimise the operation of a multi-purpose reservoir. One version used $u(t) = f(x(t))$ while the other took

the form of eqn (7.5). This form increased the computation required by a factor of three, but he found it gave only a 0.5% reduction in expected cost of operation. The problem involved correlations of up to 0.955. However, he discretised the state variable with very large increments. The smallest standard deviation for a monthly flow is 1/6th of an increment. This could invalidate some of the conclusions drawn from his work. We decided that some investigation on our problem was worthwhile.

7.4 A LINEAR QUADRATIC GAUSSIAN MODEL

To examine the effects of inflow serial correlations a linear quadratic gaussian model was formulated, in the standard form for such problems. It was attempted to make this model as much like the hydro-thermal problem as possible. An exact, analytical, solution can be obtained for the L.Q.G. problem, so the performance of the two release rules $u(t)=f(x(t))$ and $u(t)=f(x(t),I(t-1))$ can be compared easily.

The state equation is:

$$x_{i+1} = x_i + I_i - U_i \tag{7.6}$$

The cost function is:

$$J = E\left\{\frac{1}{2}(x_N-y_N)S_N(x_N-y_N) + \frac{1}{2}\sum_{i=0}^{N-1}\left|(x_i-y_i)A_i(x_i-y_i) + (d_i-u_i)B_i(d_i-u_i)\right|\right\} \tag{7.7}$$

where d_i = demand,

x_i = storage,

u_i = release,

I_i = inflow, follows equation (7.1),

y_i = (Max. Storage - Min Storage)/2,

N = final time.

A_i provides the upper and lower constraints on storage, by penalising deviations from half the maximum storage.

B_i gives the cost of thermal fuel, as a quadratic function of the thermal generation required.

Detailed model development is given in Appendix II, showing how deterministic and stochastic components are separated, etc. The optimal control law is:

$$u_i = \tilde{u}_i + \hat{u}_i = -\left(\tilde{C}_{i,1}.\tilde{x}_{i,1} + \tilde{C}_{i,2}.\tilde{x}_{i,2} + \hat{C}_{i,1}.\hat{x}_{i,1} + \hat{C}_{i,2}.\hat{x}_{i,2} \right) \qquad (7.8)$$

where $\tilde{u}_i$ = stochastic control component,

$\hat{u}_i$ = deterministic control,

$\tilde{x}_{i,1}$ = stochastic storage,

$\tilde{x}_{i,2} = I_i$ = stochastic inflow component,

$\hat{x}_{i,1}$ = deterministic storage,

$\hat{x}_{i,2} = -d_i + \hat{I}_i + \hat{T}_i$ = - demand + mean inflow + mean tributary flow,

$\hat{C}_{i,1}$ etc. are various control gains, available in closed form.

We used $B_i = 0.041$ and $y_i = 1400$ for all i, with $p_N = 1.0$ and $x(0) = 1400$. The value of A_i (a constant for all $i<N$) was chosen for each run to keep the minimum value of $\hat{x} - 1.5\sigma_x$ close to 25, (approximately 1% of maximum storage) below the actual minimum storage. Inflow and tributary data is taken from the tables in section 7.5.

A summary of the results is given in Table 7.2. The releases of the form $u_i=f(x_i, I_{i-1})$ are calculated from eqn (7.8). $u_i=f(x_i)$ results are obtained by simply setting $\tilde{C}_{i,2} = 0$ and solving with various penalty weights until the deviation from the mean state (i.e. $\hat{x}-1.5\sigma_x$) is near the desired value. Expected fuel costs increase by about 33% as inflow correlations go from 0 up to 0.6. The two control strategies give similar fuel costs, but the simpler $u_i=f(x_i)$ rule requires higher state penalties to give comparable deviations from half storage. As a result the total costs are higher for this form.

At extreme storage values, knowledge of the previous inflow would be especially useful. If a high inflow had just occurred, then a high inflow is more likely for the next time period, for non-zero serial correlation. So a larger release can now be made, for a given level of constraint violation, than if a small inflow had occurred previously.

The objective of the optimisation is to minimise fuel costs, within constraints. Both control laws $u_i=f(x_i)$ and $u_i=f(x_i, I_{i-1})$, give similar fuel costs, within the same range of permissible state variable values. The different penalty weights required are of no consequence, as they do not themselves increase fuel costs, so this test shows no advantage in using the more complicated control law.

TABLE 7.2 Comparison of Two Different Strategies for Linear Quadratic Gaussian Problem

	$u_i = f(x_i)$			$u_i = f(x_i, I_{i-1})$		
Serial Correlation Coefficient	State Penalty Weight	Fuel Cost	Total Cost	State Penalty Weight	Fuel Cost	Total Cost
0.0	.0075	68.8	100.9	.0075	68.8	100.9
0.1	.0082	71.7	107.4	.008	72.5	106.8
0.2	.009	75.1	114.9	.0082	77.0	111.7
0.3	.0099	78.8	123.3	.0083	79.8	116.4
0.4	.0108	82.6	132.4	.0084	84.0	121.2
0.5	.012	87.1	143.4	.0084	88.2	125.4
0.6	.0082	91.8	128.2	.0082	91.9	128.2

7.5 DATA FOR STOCHASTIC MODELS

All stochastic models used the same set of data to enable comparisons of their results, although inflows are in two different formats. 13 time intervals of 4 weeks each were used. The single reservoir modelled is based on the sum of all the New Zealand reservoirs, to give realistic proportions. The reservoir constraints are:

Maximum storage = $\bar{X}$ = 2850 GWh

Minimum storage = $\underline{X}$ = 0

Maximum release = $\bar{U}$ = 1500 GWh/4 weeks

Minimum release = $\underline{U}$ = 100 Gwh/4 weeks.

Thermal station data is the same as that for the deterministic model of Chapter 5. The data is reproduced here in Table 7.3, as applicable to 4 week time intervals. Thermal stations are always loaded in the order given, each station being loaded to capacity before the next most expensive is brought into use. This is the same treatment as used by the large deterministic model of Chapters 5 and 6.

Demand data is given in Table 7.4, and is considered to be known exactly.

TABLE 7.3 Thermal Station data for stochastic models

Station	Installed Capacity (MW)	Availability (%)	Generation Per 4 Weeks (GWh)	Incremental Fuel Cost ($/MWh)
New Plymouth	538	66.97	242.1	13.5
Stratford	229	75.20	115.7	16.9
Meremere	169	68.52	77.8	22.0
Marsden	233	74.71	117.0	83.4
Whirinaki	43	47.12	13.6	111.8
Otahuhu	100	57.16	38.4	120.0
Shortage	-	100.0	-	200.0

TABLE 7.4 Demand data for stochastic models. "Corrected Demand" = Actual Demand x 1.2 - Constant generation sources - mean tributary flow. This value is used for the models.

Time (4 week intervals)	Actual Demand (GWh)	Corrected Demand (GWh)
1	1343	1196.4
2	1465	1331.4
3	1708	1639.3
4	1823	1802.2
5	1729	1703.1
6	1673	1629.3
7	1586	1461.9
8	1510	1318.0
9	1446	1193.2
10	1310	1033.2
11	1254	1003.9
12	1333	1123
13	1369	1197

TABLE 7.5 Inflow data for non-linear programming stochastic models. All flows in GWh.

Time Interval	Controllable Mean	Controllable σ	Tributary Mean	Tributary σ	Total σ
1	1043.2	248.7	274.6	81.8	330.0
2	1057.4	282.7	286.6	93.5	376.3
3	1003.5	223.9	269.7	70.4	294.4
4	959.7	184.2	244.8	45.6	230.0
5	931.7	184.5	230.9	49.2	233.8
6	940.4	207.3	237.7	61.3	268.7
7	1134.9	324.2	301.1	124.8	449.0
8	1248.1	291.8	354.0	101.4	393.2
9	1328.6	260.4	401.8	101.3	361.7
10	1321.1	280.4	398.8	108.0	388.4
11	1264.5	245.6	360.1	91.4	337.0
12	1241.3	338.4	335.7	116.8	455.3
13	1118.6	303.1	306.0	109.3	412.4
Total	14,592	952	4,002	332	1,283

TABLE 7.6 Inflow data for dynamic programming models.

Probability:	0.05	0.1	0.1	0.1	0.15	0.15	0.1	0.1	0.1	0.05
Time			Inflow, Controllable, GWh per 4 weeks							
1	1679	1371	1215	1156	1064	962	910	824	756	641
2	1517	1476	1381	1196	1119	980	846	779	692	592
3	1479	1291	1198	1114	1058	945	900	747	711	657
4	1446	1210	1080	1009	962	919	864	818	738	667
5	1352	1237	1058	983	938	880	823	782	721	620
6	1503	1157	1097	1015	955	901	820	771	697	622
7	2074	1552	1343	1203	1125	1031	936	876	809	717
8	1824	1608	1500	1408	1313	1153	1030	969	907	750
9	1937	1704	1531	1457	1346	1253	1160	1087	1028	903
10	2208	1607	1474	1344	1326	1259	1182	1139	1028	916
11	1912	1586	1488	1347	1245	1173	1113	1069	1013	890
12	2175	1631	1508	1391	1244	1110	996	956	898	831
13	1829	1626	1346	1181	1110	1038	947	848	743	716
Probability:	0.05	0.1	0.1	0.1	0.15	0.15	0.1	0.1	0.1	0.05
Time			Inflow, Tributaries, GWh per 4 weeks							
1	479	397	339	289	272	249	211	217	183	176
2	477	418	389	305	305	250	214	212	179	154
3	419	341	354	297	272	271	215	197	184	172
4	371	288	266	265	249	228	223	215	211	160
5	352	303	267	246	235	199	210	198	179	160
6	403	309	246	259	246	232	222	186	167	141
7	714	441	338	322	296	256	236	217	185	176
8	553	498	444	414	371	322	260	256	258	189
9	633	537	486	436	424	361	335	331	278	244
10	710	513	444	433	389	396	361	317	271	236
11	618	498	376	361	374	346	292	281	286	238
12	639	471	435	356	322	306	269	238	222	208
13	572	471	382	332	287	273	271	206	188	167

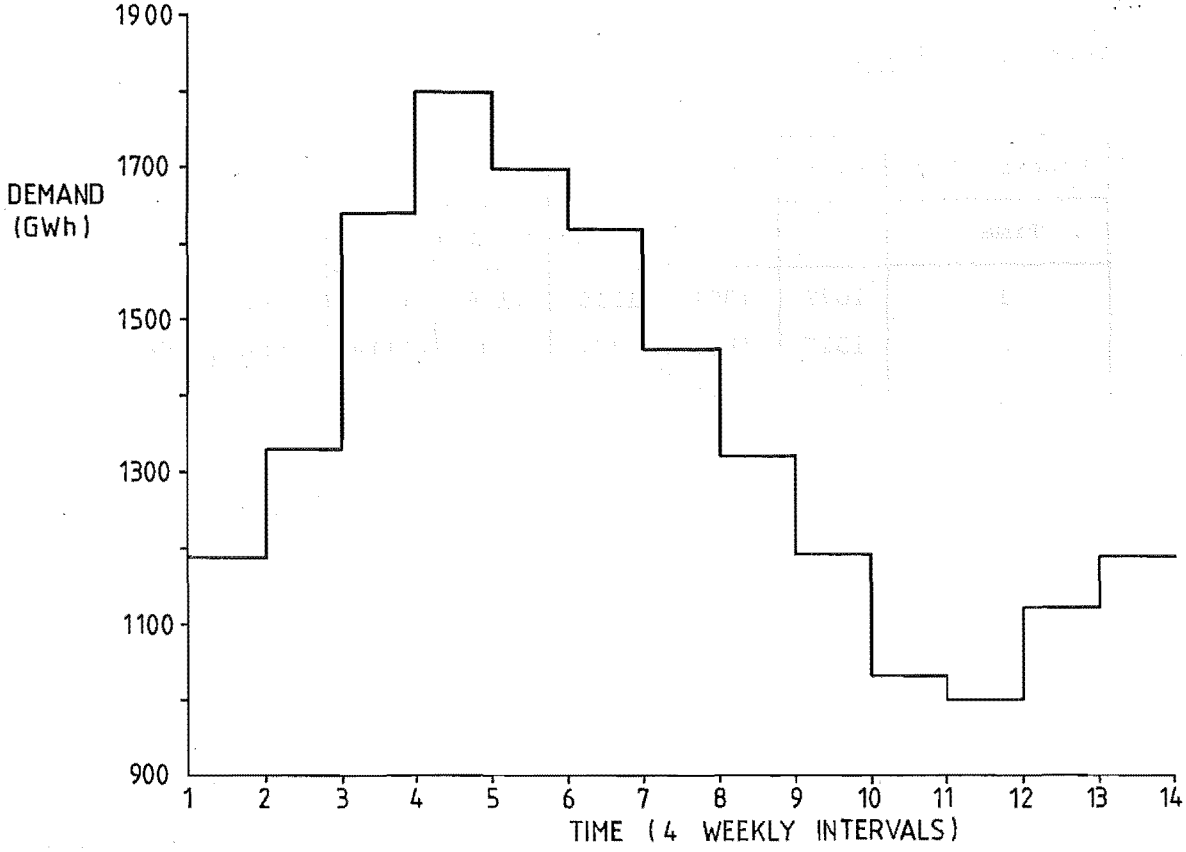


FIGURE 7.1 "Corrected Demand" pattern for one reservoir models.

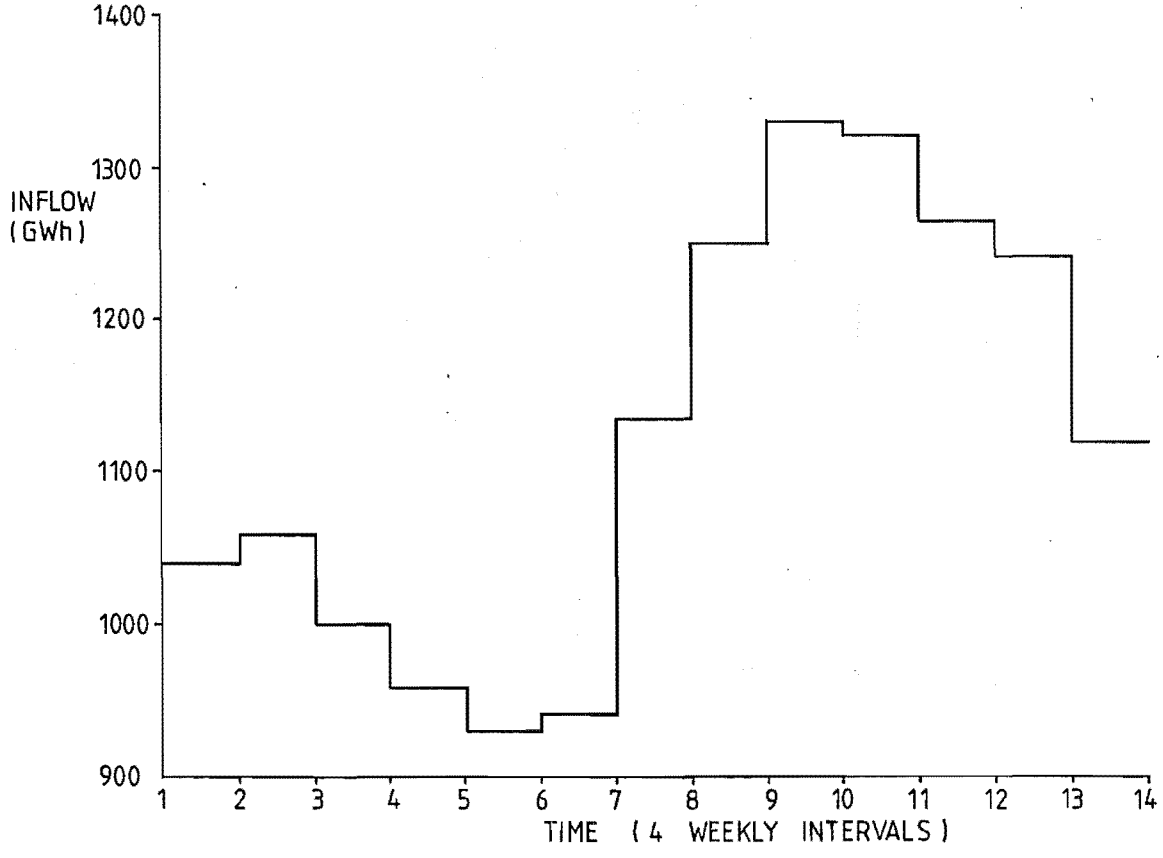


FIGURE 7.2 Inflows, total monthly averages for one reservoir models.

Inflow data is derived from the 1934-1975 period figures, part of which was used for the deterministic model. The inflows into the various reservoirs are assumed to be perfectly spatially correlated, so that they can be added together directly. Data for the non-linear programming models is given by Table 7.5. These models require flows to be treated as Gaussian, so mean and standard deviation only are needed. Separate figures are given for controllable and uncontrollable flows, although they are considered to be perfectly correlated also.

Dynamic programming models used the data of Table 7.6. Here each time interval's flow distribution is replaced by a set of 10 points. The cumulative probability distribution for each time interval has been divided into 10 probability intervals of varying widths. The average inflow for each interval has been found, and is given in the table.

7.6 STOCHASTIC AND DETERMINISTIC SOLUTIONS BY DYNAMIC PROGRAMMING

7.6.1 Methodology

Both these solutions use the same data set, from the preceding section. The aggregation of all reservoirs effectively eases storage constraints, so to compensate demands have been increased above the 1978 figures by 20%. Otherwise very little thermal generation would be needed. Also these models only consider average monthly energy needs, not peak power requirements. Storages and releases (states and controls respectively) are discretised. Initially tests were done with 50 GWh steps (28 for controls, 57 for states) but later 10 GWh steps were used (141 for controls, 286 for states). This increased computer processing time for the stochastic D.P. from 4 minutes to 55 minutes - an indication of the rapid increase in processing requirements with dimension.

Non-storable tributary flows are assumed to be perfectly correlated, in time, with storable inflows. This permits the mean value of each month's tributary inflow to be subtracted from both the demand and the 10 tributary data points. The resulting zero mean sets of data are added to the controllable inflows, increasing their spread. This means that generation is unaffected by tributary flow variation, transferring the effect to reservoir storages.

The stochastic D.P. calculation procedure is shown in Figure 7.3 for one time step. The same process is repeated for each time step, beginning at the last, working back to the first. For each state increment, at each time, only an expected "cost to go" and optimal release need be stored. The process begins with the smallest state and smallest release, testing successively larger releases until all have been tried. The minimum expected cost and corresponding release are stored. The process of testing releases then begins again for the next largest storage.

If the minimum storage constraint will be violated with a given release decision, should the minimum possible inflow occur, then that release and all possible larger releases are not feasible. No further releases need be tested for that state, and calculations move on to the next state. Testing smaller releases first thereby permits some reduction in effort. If this constraint is not violated, then the expected cost of the decision is determined using the recursive equation:

$$J_j(x_i(t)) = F(D(t)-u_j(t)) + \sum_{\ell=1}^{10} p(\ell).\ J(x_\ell(t+1)) \qquad (7.9)$$

where $x_i(t)$ is the state currently being tested,

$u_j(t)$ is the decision being tested,

$J_j(x_i(t))$ = expected "cost to go" if decision $u_j(t)$ is made when at state $x_i(t)$,

$D(t)$ = demand,

$F(.)$ = thermal fuel cost incurred by decision $u_j(t)$,

$p(\ell)$ = probability of ℓth inflow data point,

$J(x_\ell(t+1))$ = optimal expected "cost to go" from state $x_\ell(t+1)$,

and

$$x_\ell(t+1) = \text{Min}(X_{max}, x_i(t) - u_j(t) + I_\ell(t)). \qquad (7.10)$$

So the upper state constraint is enforced by simply clipping off any storage above the maximum. This is a desirable feature as it is likely to be optimal in some situations to risk spilling water with high inflows to ensure that storage is high even with lower inflows. This might be the case at the end of summer when demands have been low and inflows high, for the previous months. The strategy would ensure that the reservoir was near its maximum ready for the high demand, low inflow winter period.

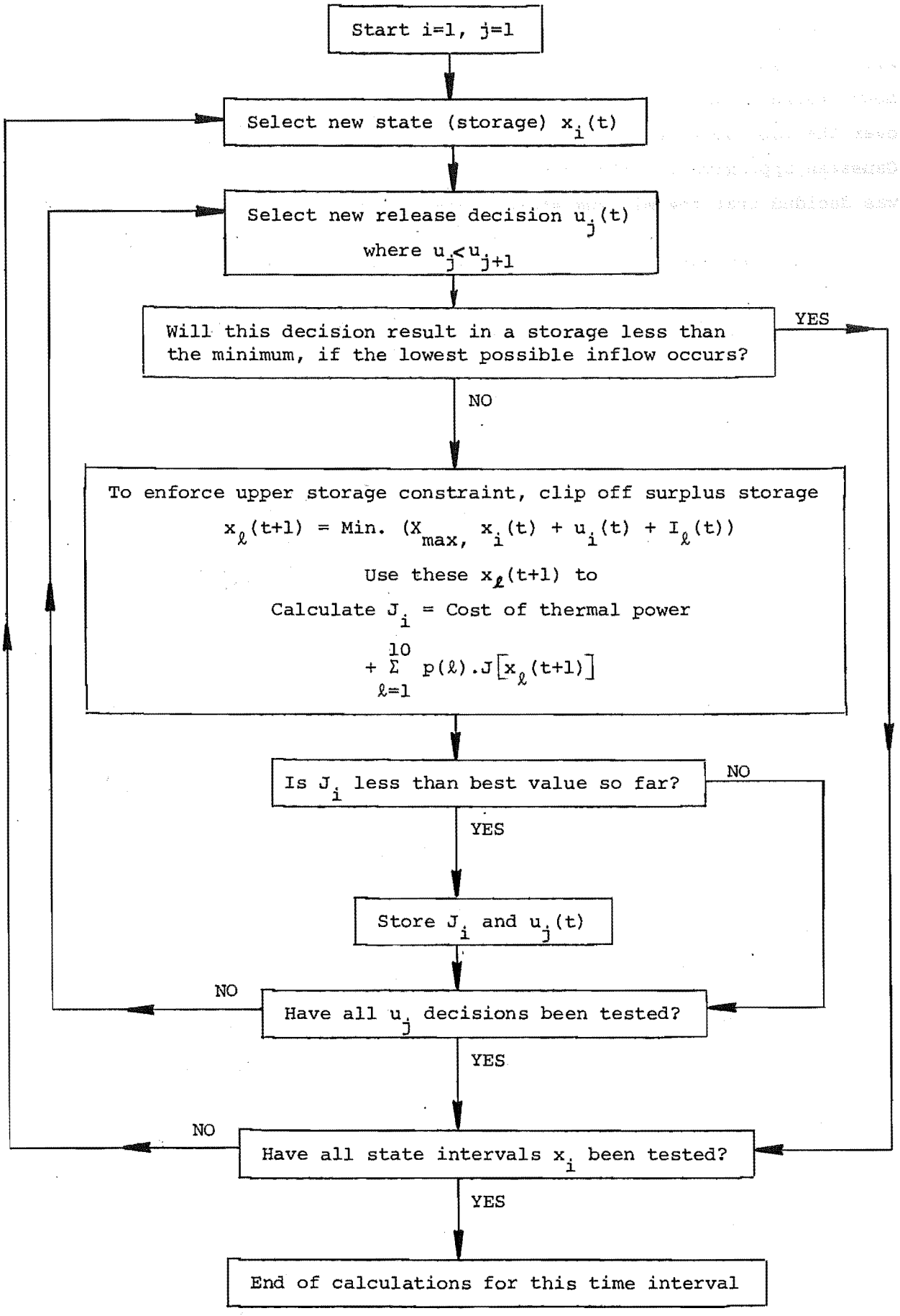


FIGURE 7.3 Flow chart of stochastic dynamic program operations for one time step.

Chance constraints are not necessary for the lower storage constraint either. Inflows are represented as ten data points, the lowest point each month representing the average value of the two lowest inflows that occurred over the last 40 years. These values correspond to about $\mu-1.5\sigma$ of the Gaussian approximation also given in the previous section. As a result it was decided that the minimum storage constraint should always be satisfied.

When beginning the solution process, some values must be supplied for $J(x_i(t_{final}))$ - the final time storage values. A target storage level can not be specified as it is not possible to be sure of reaching any particular level due to the uncertainty in the inflows. So we can not work backwards from a single storage point. A final time water value function could be used, relating the value of water in storage to its future generation potential. No obvious way exists for determining such a function. Instead, the dynamic program is repeated several times and arbitrary final time values used. The final time effects decay like a damped, transient disturbance.

On the first cycle, all final time costs are set to zero, i.e.

$$J(x_i(t_{final})) = 0 \text{ for all } x_i \tag{7.11}$$

At the end of the calculations for the year, the costs for the first time interval become the final time costs for the next cycle, after subtracting the cost for the highest storage state from all the others. This gives zero cost for maximum final time storage, and increasing costs as storage decreases. It is therefore a measure of the benefit obtained during the past year from using water, rather than retaining it for future use.

The dynamic programming gain, g_i, is defined as:

$$g_i = J(x_i(1)) - J(x_i(t_{N+1})) \tag{7.12}$$

where N time intervals equal one year.
Convergence has been reached when all gains g_i (for each state increment x_i) do not change from one cycle to the next, within some tolerance. It was found that 5 cycles were needed for a tolerance of 10, with a gain of $48{\cdot}5 \times 10^6$. All states then have the same gain.

The deterministic D.P. run to simulate O.L.F.C. performance was identical in all respects except that the 10 inflow data sets were replaced by a single averaged set.

7.6.2 Results

Output from both stochastic and deterministic D.P's is a table of releases for each time interval, for storages in 10 GWh steps. Figure 7.4 presents one form of summarising this mass of data, for the stochastic D.P. It shows how incremental fuel costs vary with storage over the year. Costs increase as storage decreases, because more thermal generation is required to offset the increased risk of running short of water, should low inflows occur. When the cost equals 200, some demand can not be met at all. Comparing the pattern with inflows and demands in Figure 7.2, the shortfalls in generation occur over the period of higher demands. No thermal generation is required for the greatest range of storages around the high inflow, low demand period.

This diagram (Figure 7.4) could be used to replace the "basic rule curve" method, mentioned in section 2.2, for ensuring secure supply. Using this diagram, the level of thermal generation needed to meet average monthly energy requirements can be found. Obviously, storage should not be allowed to get too close to the "200" zone. If it does go into this zone, then maximum thermal generation should be used, at all times. Even so, low inflows could still cause a shortfall in generation in this situation.

An interesting property of the S.D.P. is shown in Figure 7.5. No matter what storage the operation starts from, if mean flows were to occur then the storage trajectory always converges to the same steady state trajectory. This behaviour looks rather like a damped system with an impulse applied, always returning to its steady state. As the system gain converges to within less than 1 part in 10^6 after 5 cycles, this seems consistent.

Storage trajectories for stochastic and deterministic D.P's are compared in Figures 7.6 and 7.7. With perfect future knowledge, the deterministic version can go closer to the minimum storage level in Figure 7.6. The two sets of inflows simulated in Figure 7.7 are for mean $\pm \sigma_{Inflow}/3$. These sequences give annual totals of approximately the mean $\pm \sigma$, of the distribution of total annual inflows as these inflows are modelled as uncorrelated in time.

Should this particular low inflow sequence (mean - $\sigma/3$) occur, then following the specified strategy at each stage, as an O.L.F.C. would, the result is zero storage on two occasions. So it is not possible to make the specified releases at these times. For high storage levels, spilling is

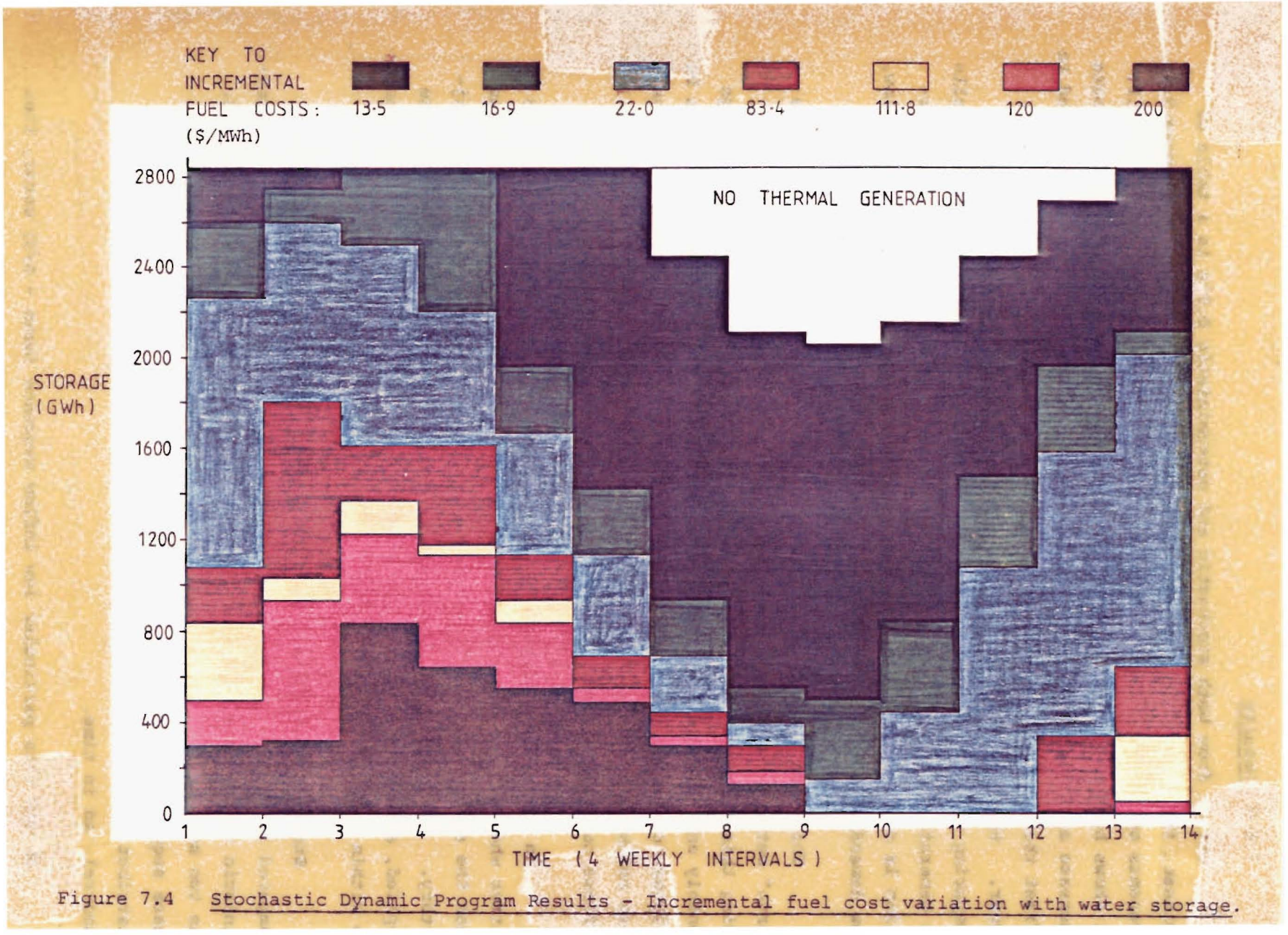


Figure 7.4 Stochastic Dynamic Program Results - Incremental fuel cost variation with water storage.

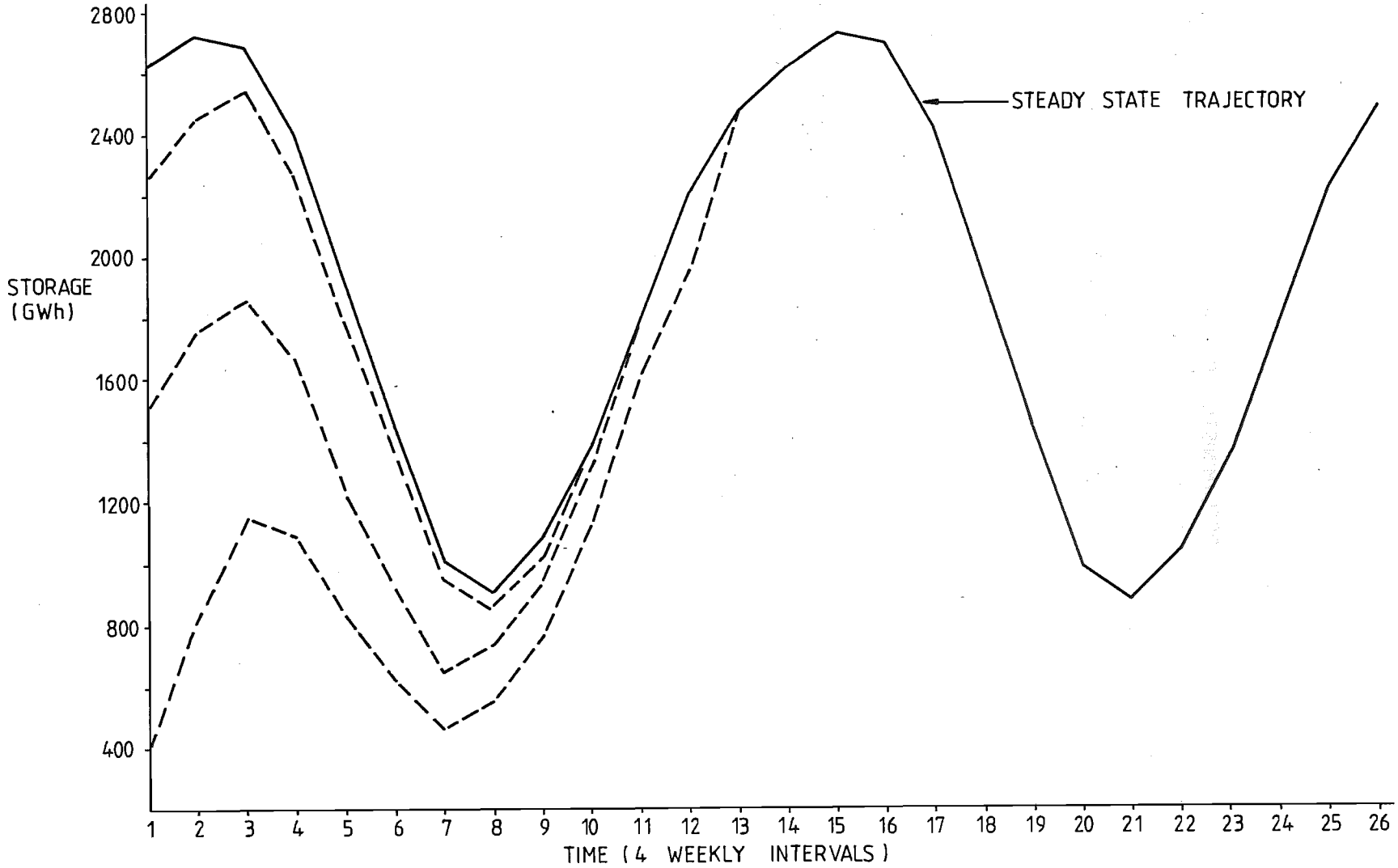


FIGURE 7.5 Mean inflows storage trajectories for stochastic dynamic program showing convergence to a steady state trajectory.

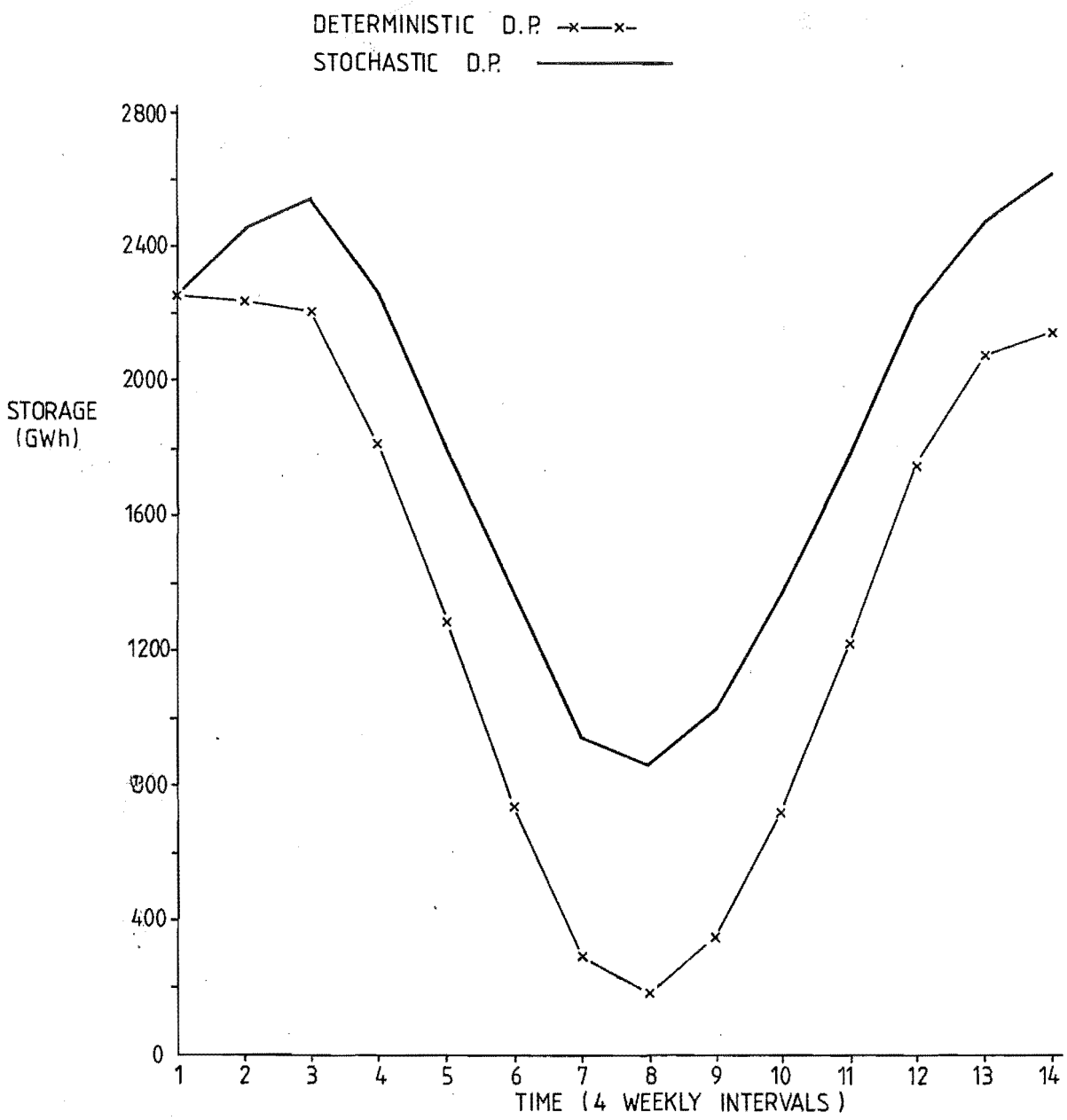


FIGURE 7.6 Mean inflows storage trajectories for stochastic and deterministic dynamic programs.

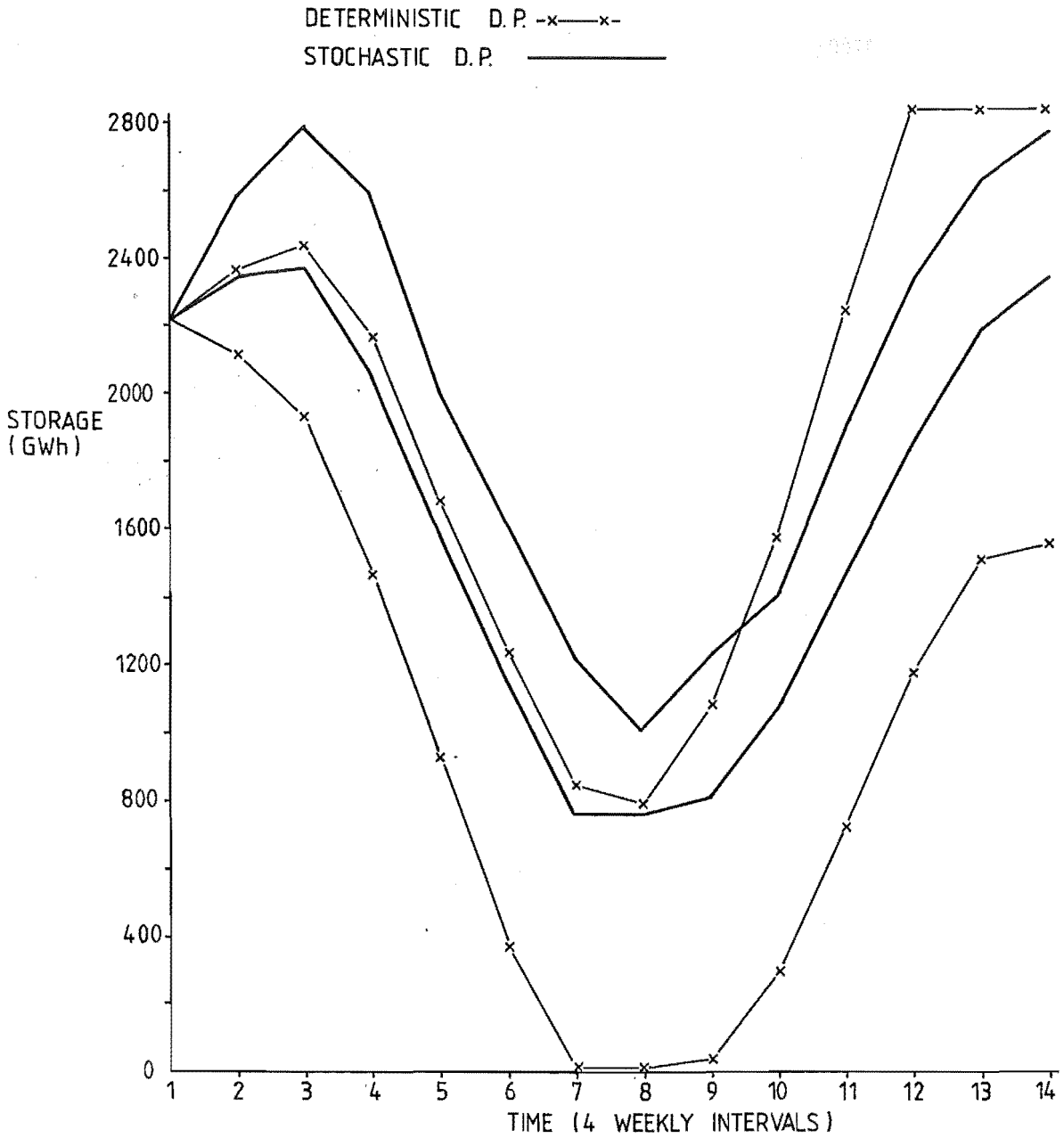


FIGURE 7.7 Mean ± σ/3 inflows storage trajectories for stochastic and deterministic dynamic programs.

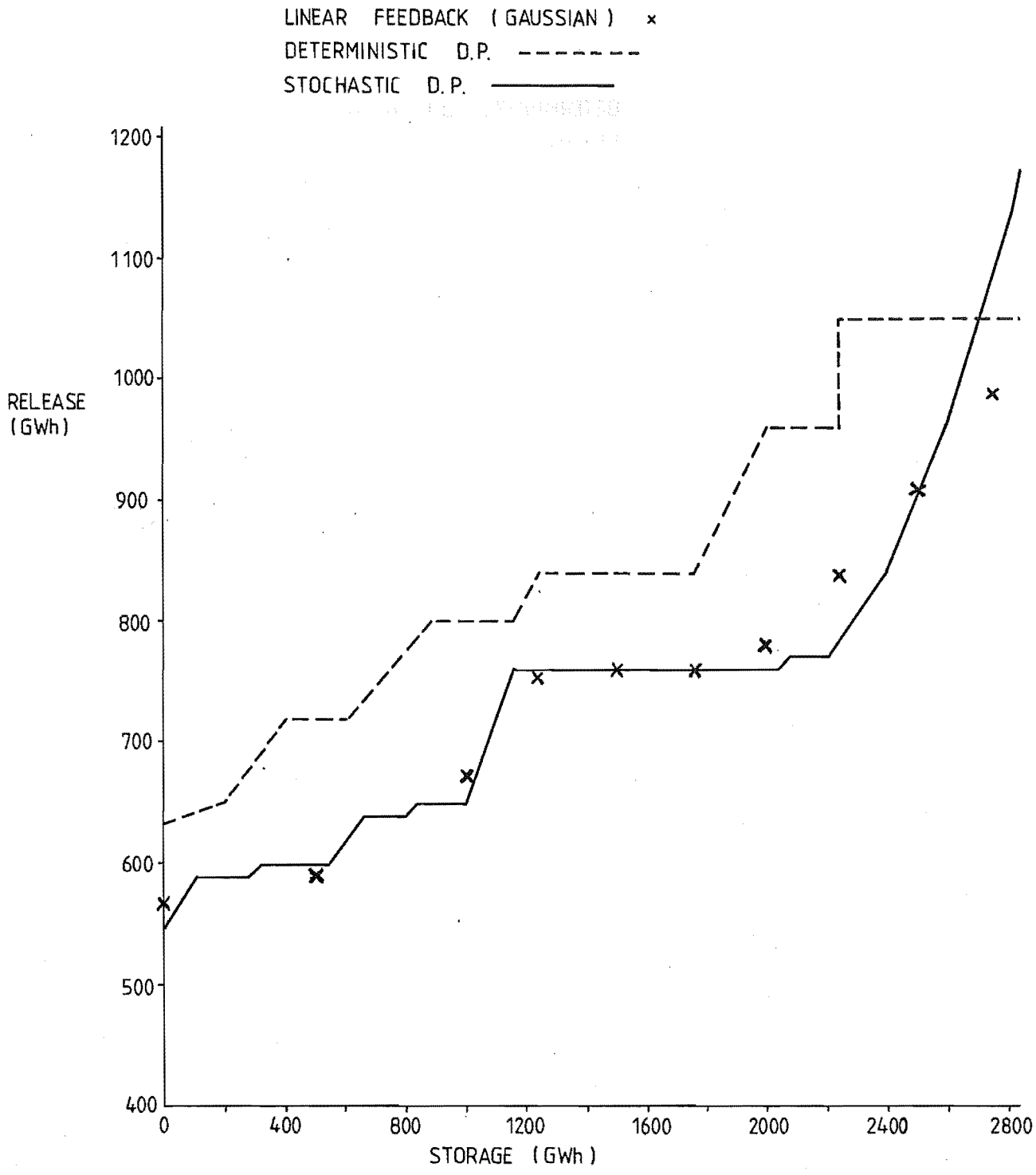


FIGURE 7.8 Releases for T=1 (first time interval).

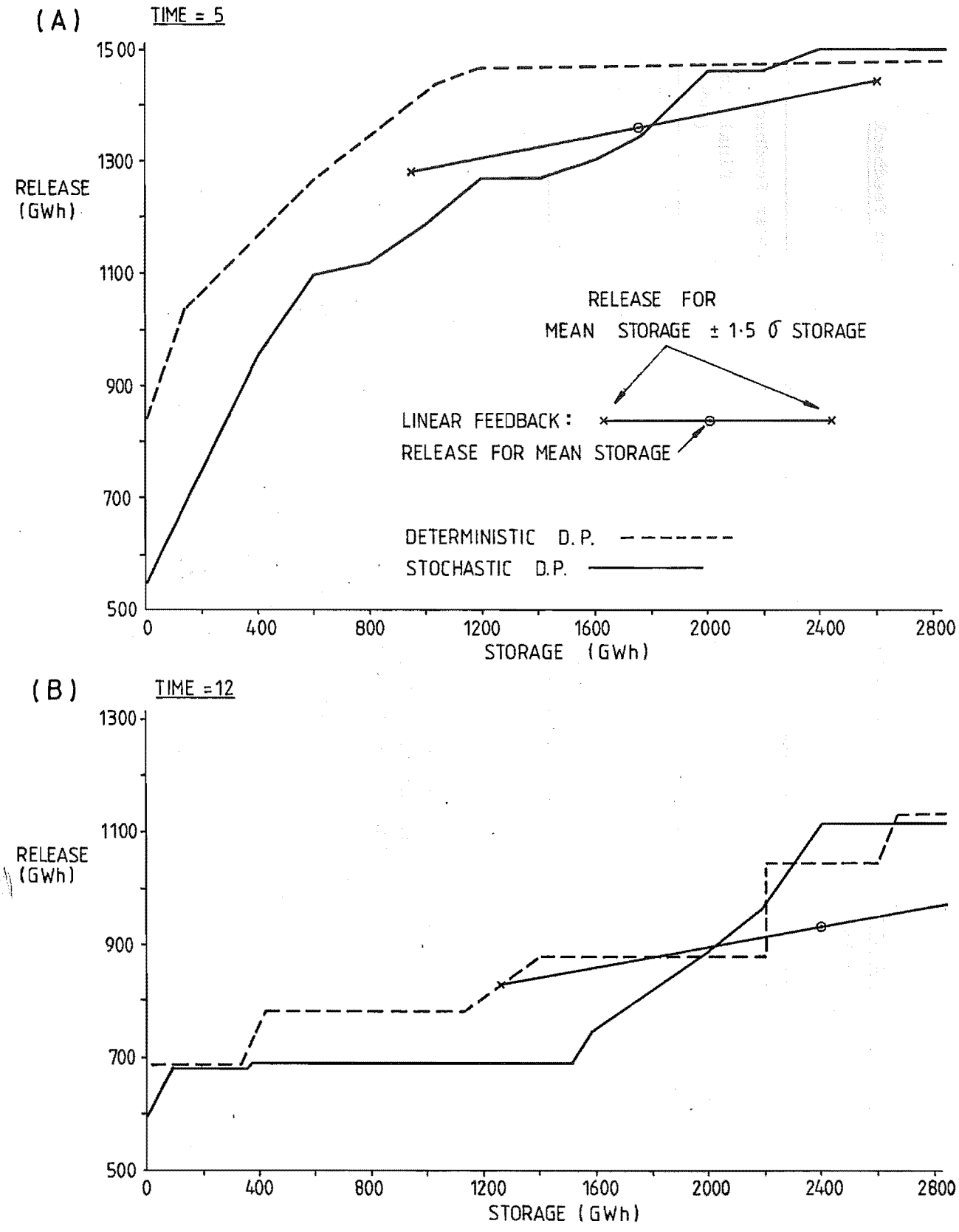


FIGURE 7.9 Releases for t=5 and t=12. Linear feedback controller results are for mean storage ±1.5σ (a) Gain = .094, (b) Gain = 0.102.

TABLE 7.7 Comparison of Costs from Stochastic Optimisations, Initial Storage = 2250, Linear Feedback Version is for 3 Cycles.

(a)

Inflows	Stockastic D.P.		Deterministic D.P.		Linear Feedback	
	Cost ($ x 10^6)	Final Storage (GWh)	Cost ($ x 10^6)	Final Storage (GWh)	Cost ($ x 10^6)	Final Storage (GWh)
$\mu - \sigma/3$	66.6	2310	56.2	1560	58.2	1985
μ	48.5	2630	43.3	2160	51.4	2828
$\mu + \sigma/3$	30.8	2770	39.2	2850	44.7	2850

(b)

Inflows	Total Annual Fuel Cost + Stochastic D.P. Final Time Storage Penalty ($ x 10^6)		
	Stochastic D.P.	Deterministic D.P.	Linear Feedback
$\mu - \sigma/3$	75.2	90.0	75.0
μ	51.5	52.6	51.7
$\mu + \sigma/3$	31.9	39.2	44.7

more likely with the deterministic D.P. Overall, the stochastic D.P. avoids both spilling and water shortages, giving a much reduced spread in storages.

Figure 7.8 indicates how this reduction in spreading is achieved - releases take a greater range of values with storage variations. This effect of increasing release with storage is especially marked for high storages, where releases increase rapidly. This is due to the high probability of spill that would otherwise apply. Two other time intervals' results are shown in Figure 7.9. In 7.9(a) the steady increase in release at the low storage end for the stochastic result is obvious.

D.P. gains obtained are $57.7 million for the stochastic, and $41.4 million for the deterministic version. These would be the expected costs in operating the system over a year, if equal initial and final storage levels were to occur. Determining the benefits of using a stochastic rather than a deterministic (O.L.F.C.) control strategy is difficult. Table 7.7(a) shows the fuel costs for the two strategies, for three inflow sequences. The deterministic is more costly only for the high inflow sequences, as it wastes some water by spilling. Final storage levels are quite different. If the final time water values of the stochastic D.P. are added to the fuel costs, in both cases, then a figure taking future water use into account accurately is obtained. The assumption is made that the S.D.P. water values are the true ones. Now the S.D.P. costs, (Table 7.7(b)), are lower in all three cases.

7.6.3 Discussion of Differences

Stochastic and open loop feedback controller (O.L.F.C.) strategies are significantly different justifying the search for a stochastic method. Some algorithm giving a water release strategy which is a function of storage is needed, but also modelling the spreading of storage - the increase in uncertainty - as time passes. The simplest means of relating storage and release is by a linear function. It can be seen from the S.D.P. plots in Figures 7.7 and 7.8 that such a function could be found, although it would be valid only over a limited range of storages. In each of these plots, three regions with linear behaviour within each can be distinguished, to varying degrees: high and low storage regions where storage constraints have a strong effect, and a middle region, less affected by constraints.

The linear feedback model of the next section was suggested by these plots.

7.7 A LINEAR FEEDBACK MODEL

7.7.1 Development

This model assumes that all stochastic variables are Gaussian and relates releases linearly to reservoir levels. The state equation is, expressing all quantities as their electrical equivalents:

$$x(t+1) = x(t) + i(t) - \bar{u}(t) - k(t)(x(t) - \bar{x}(t)) \tag{7.13}$$

with $x(1)$ given, $i(t)$ known.

where x and i are Gaussian,

$x(t)$ = storage,

$i(t)$ = inflow,

$\bar{u}(t)$ = release at mean storage, a control variable,

$k(t)$ = gain, also a control variable,

$\bar{x}(t)$ = mean storage.

The term $k(t)(x(t)-\bar{x}(t))$ adjusts the releases in proportion to the deviation of the actual Gaussian distributed storage, $x(t)$, from the mean value $\bar{x}(t)$. As the system is linear and Gaussian, the state equation can be split into mean and variance equations, which involve only deterministic quantities:

$$\bar{x}(t+1) = \bar{x}(t) + \bar{i}(t) - \bar{u}(t) \tag{7.14}$$

$$\sigma_x^2(t+1) = \sigma_i^2(t) + (1-k(t))^2\, \sigma_x^2(t) \tag{7.15}$$

The control action is given by:

$$u(t) = \bar{u}(t) + k(t)(x(t)-\bar{x}(t)) \tag{7.16}$$

where $\sigma_x^2(t)$ = storage variance,

$\sigma_i^2(t)$ = inflow variance,

$u(t)$ = release, Gaussian distributed.

Now the state variables $\bar{x}$ and σ_x^2 describe the physically observable quantity x. The control variables $\bar{u}$ and k define the physical release u.

Note that if $k(t)=1$, then $\sigma_x^2(t+1) = \sigma_i^2(t+1)$. In this case the uncertainty in $x(t+1)$ is due only to the uncertainty in the preceding period's inflow. In general, increasing $k(t)$ reduces the rate of increase in uncertainty in storage.

Constraints on the variables will be considered now. Chance constraints are applied to the minimum values of storage and release. They are enforced by penalty functions, adding the following quantities to the cost function:

$$W_x(X_{min} - (\bar{x}(t) - a\,\sigma_x(t)))^2 \text{ if } \bar{x}(t) - a\,\sigma_x(t) < X_{min}$$
$$W_u(U_{min} - (\bar{u}(t) - a\,k(t)\sigma_x(t)))^2 \qquad (7.17)$$
$$\text{if } \bar{u} - a\,k(t)\sigma_x(t) < U_{min}$$

where a determines the probability of constraint violation that is considered acceptable.

Upper constraints are more difficult. Spill is required under some circumstances, so applying a chance constraint to the upper limit of storage could be unduly restrictive. Storages could be confined to a narrow range decreasing as time goes into the future, as uncertainty increases. The permitted storage range for the system modelled is 2850 GWh, but σ=1283, for the total annual inflows (Table 7.5). So for $k(t)=0, \forall t$, a feasible solution will not exist if a = 1.5 in eqns (7.17), i.e. if the 1.5σ level of storages is required to satisfy upper and lower constraints.

Unfortunately, simply clipping storages off if they exceed their maximum introduces a non-linearity, so the system state is no longer Gaussian. The convenient separation of mean and variance components, eqns (7.14) and (7.15), would then not be possible. As a simplification, spilling is considered to occur only when the mean storage exceeds the maximum. Hence eqn (7.14) becomes:

$$\bar{x}(t+1) = \bar{x}(t) + \bar{i}(t) - \bar{u}(t) - s(t) \qquad (7.18)$$

where $s(t) = \text{spill} = \begin{cases} 0 \text{ if } \bar{x}(t)+\bar{i}(t)-\bar{u}(t) < X_{max} \\ -X_{max} + (\bar{x}(t)+\bar{i}(t)-\bar{u}(t)) \text{ otherwise.} \end{cases}$

Releases are treated in a similar way. It may be worthwhile accepting a small risk of releasing more water than can be put through turbines, with high storage levels, to ensure that generation is near the maximum for some lower, higher probability storage. To allow this, releases are not constrained at their upper limit, but any release above the limit has no more effect on thermal fuel costs than is possible at the maximum. So when calculating fuel costs, u(t) is replaced by u'(t):

$$u'(t) = \begin{cases} \bar{u}(t) + k(t)\ (x(t)-\bar{x}(t)) & \text{if } u(t) < U_{max} \\ U_{max} \text{ if } u(t) > U_{max} \end{cases} \tag{7.19}$$

Wasting water is undesirable, increasing fuel costs unnecessarily. This constrains releases sufficiently.

All that remains now is for the cost function to be defined, which is to be minimised by the optimisation algorithm. This is the cost of thermal power, plus the penalties of eqns (7.17). Thermal costs are determined in the same way as for the D.P. models, for a given release. However releases now follow a continuous distribution, so:

$$\text{Fuel Cost} = \int_{-\infty}^{\infty} F\left[D(t) - \left\{\bar{u}(t) + k(t)(x(t)-\bar{x}(t))\right\}\right].p(x(t))dx(t) \tag{7.20}$$

where $F[.]$ = cost of thermal generation for a specific release.

An attempt was made to perform this integration using a 12 point Gaussian quadrature formula, (see Stroud and Secrest, 1966). Difficulties were experienced with the non-smooth contours produced by the quadrature formulas, so error functions were then tried, successfully.

From Abramowitz and Stegun:

$$\text{erf}(x) = 1. - \frac{1}{(1+a_1 x+a_2x^2 + a_3x^3 + a_4x^4)^4} + \varepsilon(x) \tag{7.22}$$

where $|\varepsilon(x)| \leq 5 \times 10^{-4}$

and $a_1 = .278393$, $a_2 = .230389$, $a_3 = .000972$, $a_4 = 078108$. This gives areas under the normal distribution with mean m and standard deviation σ, from:

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(t-m)^2/2c^2}\, dt = \frac{1}{2}\left(1+\text{erf}\left(\frac{x-m}{\sigma\sqrt{2}}\right)\right) \tag{7.23}$$

As the thermal cost function consists of a series of linear segments it can be written as:

$$F\left[D(t) - \bar{u}(t) + k(t)(x(t) - \bar{x}(t))\right] \tag{7.24}$$

$$= A(\nu) + B(\nu)(D(t) - \bar{u}(t) - k(t)(x(t)-\bar{x}(t)))$$

for the appropriate $A(\nu)$, $B(\nu)$ which apply when the storage $x(t)$ is between $Z(\nu,t)$ and $Z(\nu+1,t)$.

$z(\nu,t)$ is the storage level below which thermal station ν is not required, and:

$$z(7,t) = \infty \qquad z(0,t) = -00 \quad \forall\ t.$$

$z(\nu,t)$ is of course dependent on demand and release.

Therefore the expected cost of eqn (7.24) becomes:

$$E[F(.)] = \sum_{\nu=0}^{7} \int_{\infty}^{\infty} \frac{1}{\sigma_x(t)\sqrt{2\pi}} e^{-(x(t)-\bar{x}(t))^2/2\sigma_x(t)^2}$$

$$\cdot \left\{A(\nu) + B(\nu)(D(t)-\bar{u}(t) - k(t)(x(t)-\bar{x}(t)))\right\} \tag{7.25}$$

$$= \frac{k(t)}{|k(t)|} \sum_{\nu=0}^{7} \frac{1}{2} \left(A(\nu) + B(\nu)(D(t)-\bar{u}(t))\left(\operatorname{erf}\left[\frac{Z(\nu,t)-\bar{x}(t)}{\sqrt{2}\,\sigma_x(t)}\right]\right.\right.$$

$$\left.\left. - \operatorname{erf}\left[\frac{Z(\nu+1,t)-\bar{x}(t)}{\sqrt{2}\,\sigma_x(t)}\right]\right) - k(t)B(\nu).\ \frac{\sigma_x(t)}{\sqrt{2\pi}}\right.$$

$$\left. \cdot \left[-e^{-(z(\nu,t)-\bar{x}(t))^2/2\sigma_x(t)^2} + e^{-(z(\nu+1,t)-\bar{x}(t))^2/2\nu_x(t)^2}\right)\right] \tag{7.26}$$

The problem so far is:

Minimise the cost of thermal power (7.26) plus the chance constraint penalty terms (7.17) subject to the state equations (7.15) and (7.18) with the controls defined by (7.16), and the restriction (7.19) when calculating the cost of thermal station fuel. Data needed is given in section 7.5, and the initial storages are specified. This problem is solved using CGRADS, but calculating gradients by a finite differences technique, described in Chapter 8. Calculation of gradients analytically is extremely complicated, so this shortcut was taken to speed development. The final aspect to be considered is the choice of final time storage target.

As for the stochastic dynamic program, a specific target storage can not be met due to the uncertainty in storage. The mean final storage could be required to be the same as the initial, but there is no reason why this should be optimal over a period of several years. Instead, a similar approach to the stochastic D.P. is used. The final time mean storage is required to be greater then some arbitrary value. This is enforced by a penalty function. The optimisation is then performed for several years, not just one. When the final time is sufficiently far into the future it

should not affect the first time period's solution, which is the only one that will actually be used. (The algorithm would be used as an O.L.F.C.). Like the stochastic D.P., it is hoped that the disturbance created by an arbitrary target mean storage will die out, given sufficient time.

7.7.2 Results

Convergence of the algorithm is discussed first, followed by a comparison with stochastic D.P. results for storage trajectories, initial time release pattern, release patterns for other times, and then of costs for various inflow levels.

Table 7.8 shows releases for the first time interval for various runs. If satisfactorily converged, the final storage value will not affect the release for t=1 - its effect will have died out. Hence the comparison of pairs of results for final storages of 2250 GWh and 1690 GWh, (a 25% reduction). None of the results listed were affected greatly by the final storage, even for only one cycle. The largest change (for the two runs of 2 cycles, initial storage = 2250 GWh) is 0.98%. A change of only 0.22% occurs for the three cycle runs, which seems fairly satisfactory.

TABLE 7.8 Releases for t=1 for Various Runs of Linear Feedback Model

Number of Cycles	Initial Storage = 2250 Mean Final Storage: 2250	1690	Initial Storage = 1500 Mean Final Storage: 2250	1690
1	838.6	838.6	760.8	760.8
2	829.3	837.4	760.8	760.7
3	837.6	835.8	760.6	760.8
4	837.8	842.8	761.2	760.8

Once end effects have died out, adding another cycle should not produce any changes. However, a 1.0% difference is found in initial release decisions, between 2 and 4 cycle solutions. This is for initial storage = 2250 GWh, but practically no change occurred when initial storages of 1500 GWh were used. Repeatability was checked by trying different initial guesses for the release strategy, and trying different linear search accuracies for CGRADS . No significant changes were produced. It appears, therefore, that there must be some slight instability in the algorithm, perhaps only occurring with high initial storages.

The comparison of the mean inflows storage trajectory with the S.D.P. result, (Figure 7.10) is more encouraging. The improvement over Figure 7.6 of the O.L.F.C. result is very great. The two trajectories (true S.D.P. solution, and linear feedback approximation) are now almost indistinguishable until t=6. A similar result is shown in Figure 7.11 for mean $\pm$ $\sigma/3$ inflows. The two trajectories for the linear feedback model compare much more favourably with the S.D.P. result than does the O.L.F.C. in Figure 7.7. While the O.L.F.C. low inflows trajectory empties the reservoir completely, the chance constraints of the linear feedback controller ensure a safe margin.

The spread in storages (difference between high and low inflow trajectories) is much reduced by the feedback action. Final time spreads shown in Figures 7.11 and 7.7 are 1290 GWh for O.L.F.C., 865 GWh for linear feedback, and 460 GWh for S.D.P. Larger gains give a greater reduction in the spreading. For the 3 cycle run with initial storage of 2250 GWh, most gains were in the region of 0.1, and the maximum was 0.21.

The high inflow trajectory for the linear feedback model goes above maximum storage, according to the optimisation output. It has been drawn here as it would happen in practice - clipped off at maximum storage.

Figure 7.8 shows the initial release decision, for various initial storages. The D.P. solution methods provide all the necessary information for their plots from a single solution run, but separate linear feedback solutions are needed to obtain each point shown for this method. In this case, the initial storage is a specified input and only a single release decision need be made, with zero gain, for the first time interval. This is because there is no uncertainty in storage for t=1. Agreement with the S.D.P. plot is excellent, and the improvement on O.L.F.C. results considerable. At high storages, S.D.P. releases are larger, because the linear feedback controller has extra storage of 1.5 σ_x, effectively, as spilling does not occur until the mean storage exceeds the maximum.

Behaviour for two time periods in the future is shown in Figure 7.9, for all three algorithms. For t=5, S.D.P. and linear feedback results agree reasonably well. The S.D.P. release versus storage plot is approximately linear for some distance either side of the linear feedback controller's mean release. For t=12 the comparison is not so good. Here this mean release is near a corner in the S.D.P. plot - it switches from one linear segment to another. At t=5, demand exceeds the maximum possible

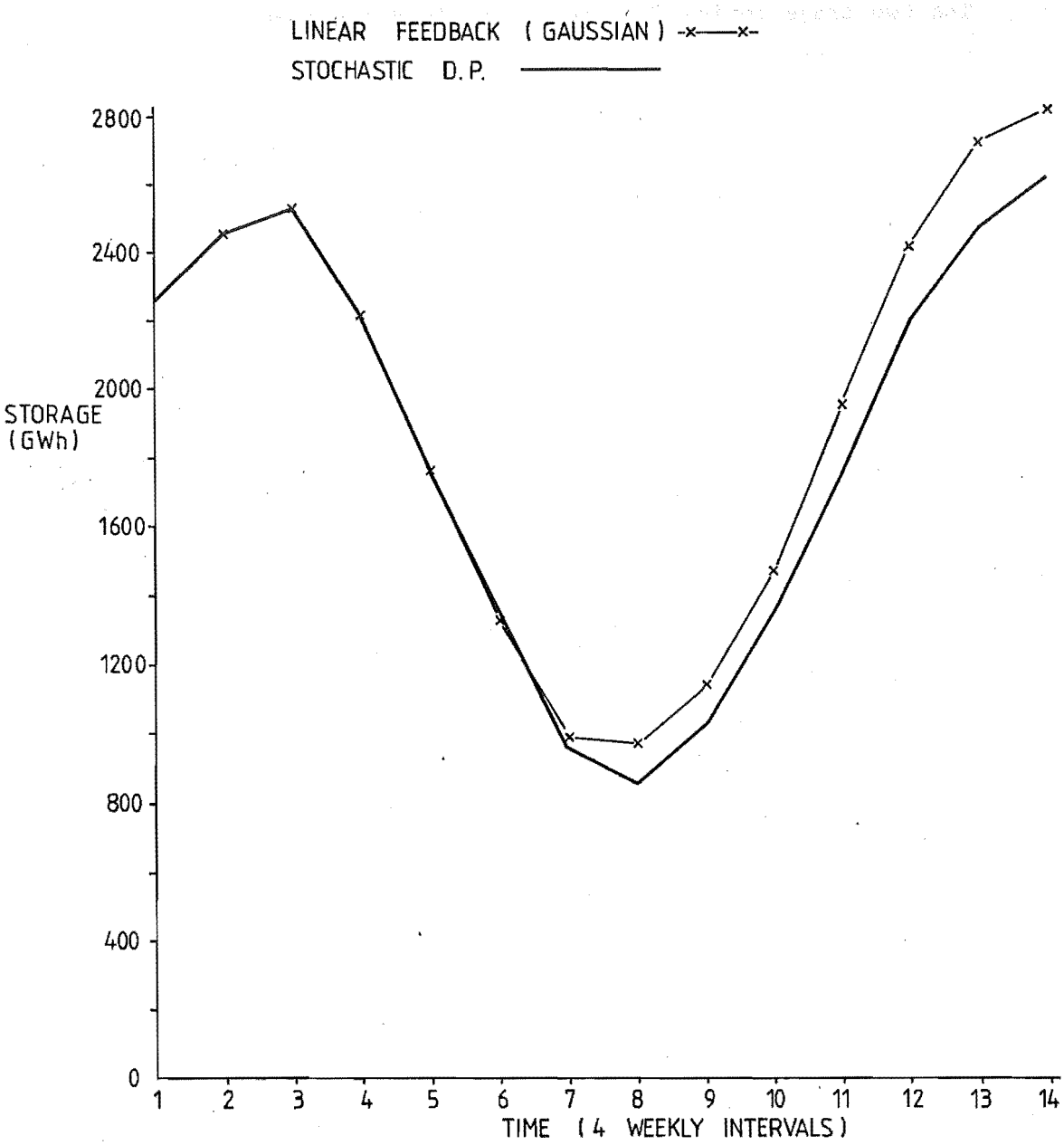


FIGURE 7.10 Mean inflows storage trajectories for stochastic dynamic program and linear feedback model with Gaussian storage distribution.

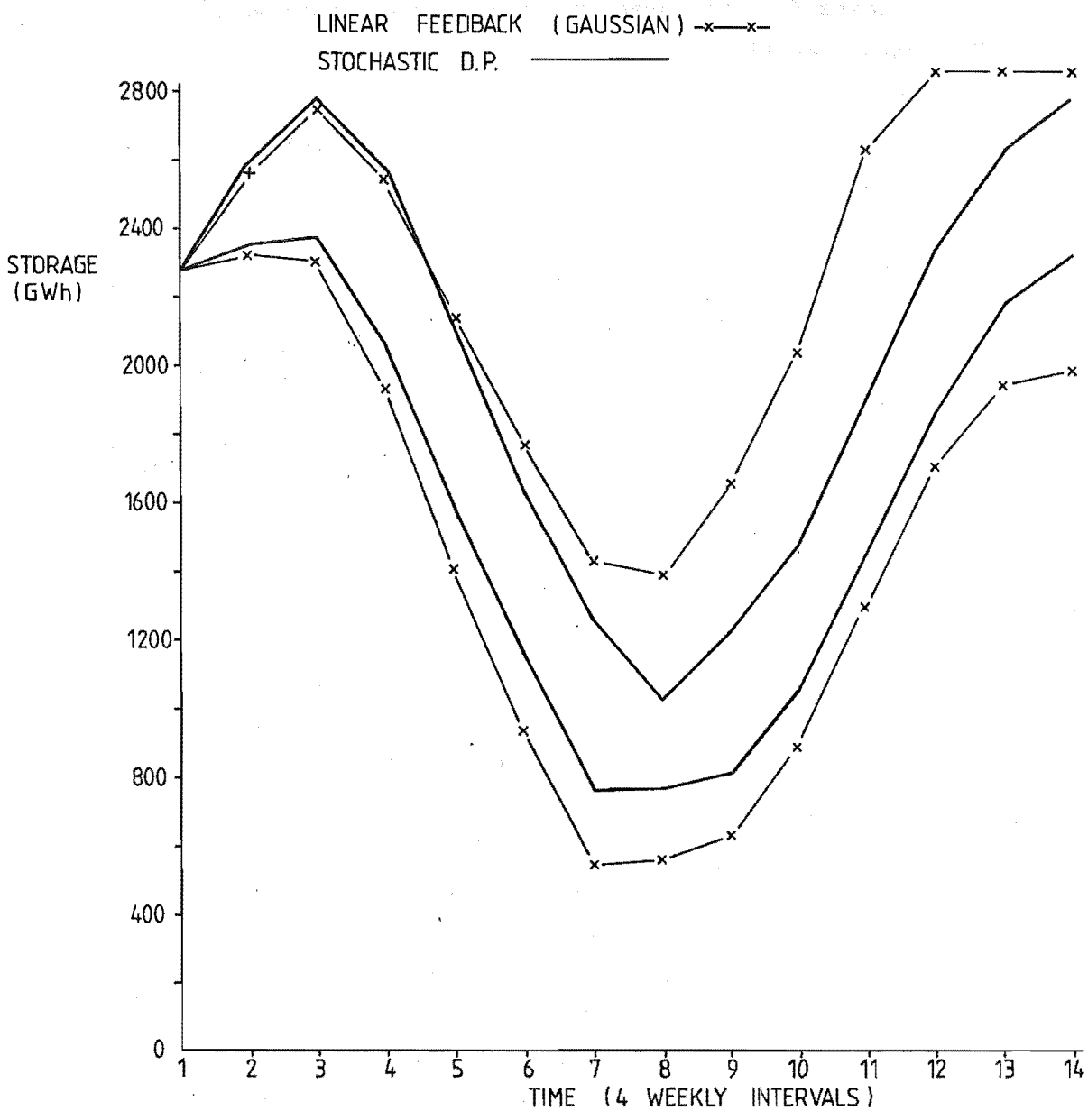


FIGURE 7.11 Mean ± σ/3 inflows storage trajectories for stochastic dynamic program and linear feedback model, with Gaussian distributed storages (3 cycles).

release (of 1500 GWh) but for t=12, demand = 1123 GWh and so for high storages the S.D.P. strategy needs no thermal generation. Hence the zero slope segment occurs when release equals demand. A linear feedback strategy with different gains for storages above and below the mean might enable this sort of behaviour to be imitated.

Costs for the three inflow sequences examined are given in Table 7.7. The costs when final storage is penalised according to the values given by the S.D.P. are most interesting. Linear feedback and S.D.P. costs are then almost identical for mean flows and the low inflows. This suggests that the lower storage constraints are approximately as severe in both cases. The two trajectories were seen to be similar in Figure 7.11, also. For high inflows, the linear feedback cost is almost $13 million higher. This is not surprising as some of the extra inflows were lost when converting the optimisation's output to the physically realisable form of Figure 7.11, i.e. when clipping off storages at their maximum.

The principle weakness of the new model is obviously in its failure to treat spilling correctly, and motivates the development of the model of Chapter 8.

7.8 CONCLUSIONS

The investigations of the stochastic hydro-thermal problem began by looking at the effect of serial correlations in inflows with a simple LQG model. They were found to have a significant effect on constraint violations, but not on fuel costs. Hence serial correlations were not modelled .

Dynamic programming showed how an O.L.F.C. differs considerably from the true stochastic D.P. solution. A sub-optimal linear feedback method, modelling storage distributions as Gaussian, was compared with the true optimal S.D.P. solution. The new method is feasible for multi-reservoir problems. It relies on the linear state (water storage) difference equation to permit separation of mean and variable (stochastic) components.

The linear feedback solutions approximated the S.D.P. results remarkably well, but problems exist with the unrealistic method of modelling reservoir spill, when maximum storage is reached. In reality this distorts the otherwise Gaussian behaviour of storage probability distributions. Chapter 8 presents an attempt to overcome this deficiency. The realistic

modelling of spill sets this hydro-thermal problem apart from other water resources type problems which can use simple methods for enforcing both upper and lower storage constraints.

CHAPTER 8

A POTENTIAL STOCHASTIC MULTI-RESERVOIR SOLUTION METHOD

8.1 INTRODUCTION

The new model described in this chapter is a development of that in the previous chapter which represented the probability distribution of reservoir storages as a Gaussian. While some comparisons with the stochastic D.P. solutions were very favourable, it was concluded that the inability to model spilling of excess storage until the probability of spill was 0.5 was a severe weakness. Now an attempt to overcome this difficulty is made, modelling the storage distribution by a Gaussian plus an impulse of probability at maximum storage.

The principle of this approach is described in section 8.2.1, followed by detailed model development (8.2.2) and results (section 8.3). As many aspects of the model as possible are identical to the Gaussian linear feedback version, but a brief outline of the one reservoir model follows.

The one reservoir has storage and release limits approximating the sum of the values used for the eight reservoir New Zealand system model. Hence the proportions are realistic. Demand data are the 1978 values, times 1.2, minus mean tributary flows. Inflow data are from the period 1934-1975. Controllable inflows have added to them the tributary variations (i.e. tributary flow minus mean tributary flow). Inflows are assumed to be Gaussian. Thermal stations accurately represent the six New Zealand stations. No transmission losses are included.

8.2 THE GAUSSIAN PLUS IMPULSE MODEL

8.2.1 The Concept - A Simple Form of Gaussian Sum

Figure 8.1 shows how the probability distribution of storage in a reservoir might change with time. The initial storage is known with certainty. If the mean storage value rises, then there is as in Figure 8.1(c) some probability, α, that maximum storage is exceeded. In reality, any excess above the maximum is spilled, so there is actually a probability α of being at the maximum. This gives rise to the impulse, with an area of α at

the maximum in (d), and the truncated Gaussian. The approximation process of the Gaussian plus impulse models this accumulation of probability at the maximum.

An alternative approach considered involved using Gaussian sums (Sorenson and Alspach 1971, Alspach and Sorenson 1972, Anderson and Moore 1979). In the same way that a function can be represented by a Fourier series, any probability distribution, $p(x)$, can be replaced by a Gaussian sum approximation, $p_n(x)$, defined as:

$$p_n(x) = \sum_{i=1}^{n} \alpha_i N_i (x-\mu_i, \sigma_i) \tag{8.1}$$

where

$$\sum_{i=1}^{n} \alpha_i = 1 \qquad \alpha_i \geq 0 \text{ for all } i.$$

A calculation procedure suggested by Sorenson and Alspach (1971) is as follows, for a uniform distribution on the interval (-2,2).

(i) Space the mean values of the Gaussians, μ_i, equally along the interval,

(ii) Set $\alpha_i = 1/n$, where n Gaussians are to be used for the approximation.

(iii) All σ_i^2 are made the same, and are selected by a one dimensional search to minimise

$$L' = ||p-p_k|| \tag{8.2}$$

where p is the true distribution, p_k the approximation.

If used with our reservoir scheduling algorithm, the reservoir storage p.d.f. would be approximated in this way. A new set of approximations would have to be found each time a function evaluation was required by CGRADS, the conjugate gradients algorithm. Small errors in the approximation, if inconsistent, could result in awkward contours in the cost function. The possibility of spurious local minima caused in this way, and the computational effort likely, resulted in the choice of a very elementary Gaussian sum approximation strategy - the "Gaussian plus impulse" method.

If the Gaussian sum approximations could be made with a sufficiently small amount of computation, and consistently, then they could offer a practical solution method for this type of problem.

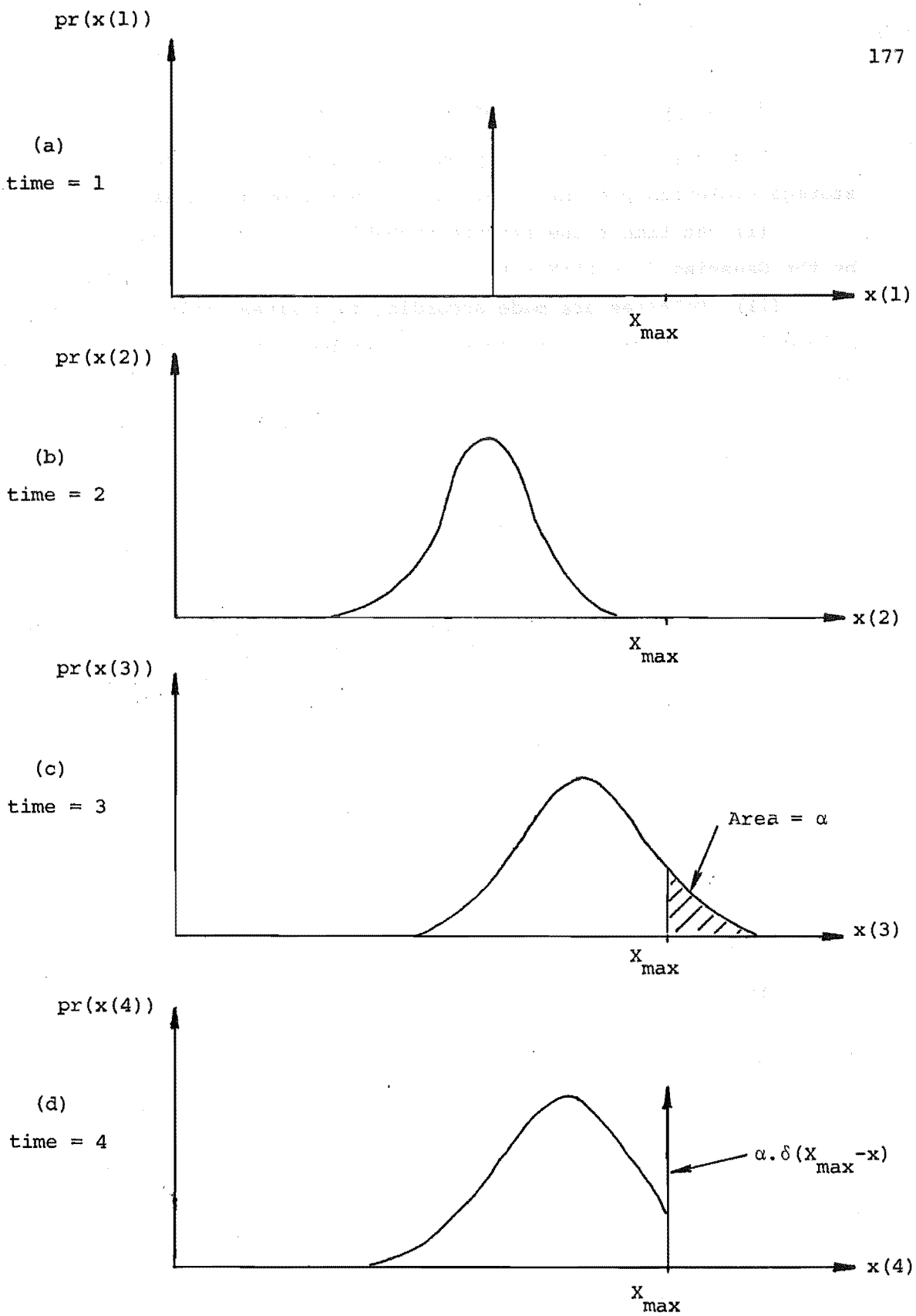


FIGURE 8.1 Accumulation of uncertainty in storage, beginning from a known storage at time 1.

(c) finite probability of spill

(d) real situation - probability of being at X_{max} increases due to spill truncating the distribution.

8.2.2 Model Development

Figure 8.2 shows the approximation process for the new model. The storage evolution process occurring at each time interval is as follows:

(i) At time t the storage probability distribution is represented by the Gaussian $(1-\alpha(t))N(\mu(t),\sigma(t))$ and the impulse $\alpha(t)\cdot\delta(X_{max}-x(t))$.

(ii) Releases are made according to a linear rule, i.e. release is a function of storage. A Gaussian distributed inflow, $N(\bar{i}(t), \sigma_i(t))$, occurs.

(iii) At time t+1 the events of (ii) have resulted in two Gaussians representing storage probability. $(1-\alpha(t))\cdot N(\mu'(t),\sigma'(t))$ and $\alpha(t)\cdot N(\mu''(t), \sigma''(t))$

(iv) First step of the approximation - determine the probability of exceeding maximum storage. This gives $\alpha(t+1)$.

(v) Second step of the approximation - from the two truncated Gaussians remaining, find a single Gaussian approximation, $(1-\alpha(t+1))N(\mu(t+1, \alpha(t+1))$.

A detailed development follows:

At time t the storage x(t) has the probability distribution:

$$p(x(t)) = (1-\alpha(t))N(\mu(t),\sigma(t)) + \alpha(t)\delta(X_{max}-x(t)) \tag{8.3}$$

where $\alpha(t)$ = probability of being at maximum storage,

X_{max} = maximum storage,

$\mu(t)$ = mean of Gaussian,

$\sigma(t)$ = standard deviation of Gaussian.

The mean storage $\bar{x}(t)$:

$$\bar{x}(t) = (1-\alpha(t))\mu(t) + \alpha(t)X_{max} \tag{8.4}$$

Define the probability transition over the period t to t+1,

$$p(x(t+1)) = (1-\alpha(t))N(\mu'(t),\sigma'(t)) + \alpha(t)N(\mu''(t),\sigma''(t))$$

$$\simeq (1-\alpha(t+1))N(\mu(t+1),\sigma(t+1)) + \alpha(t+1)\delta(X_{max}-x(t+1)) \tag{8.5}$$

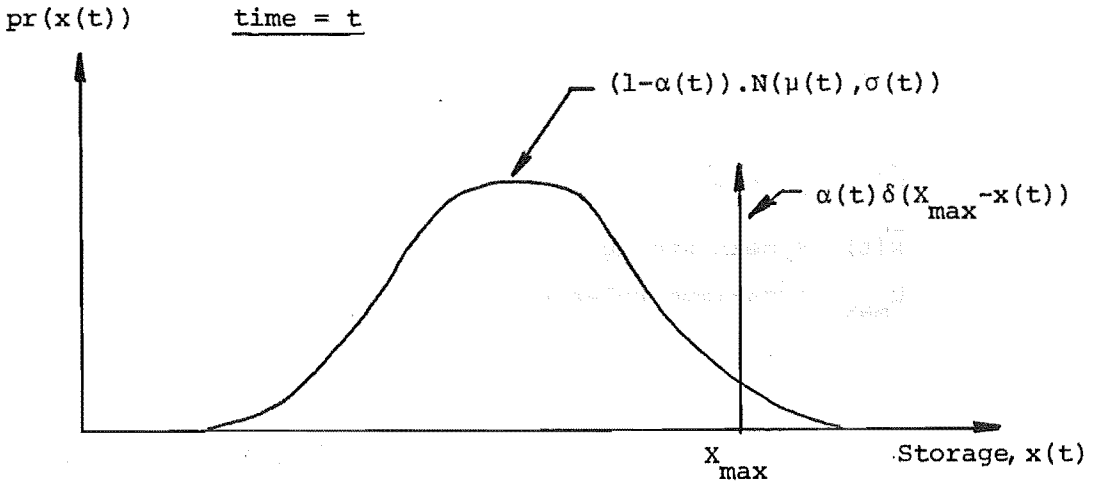
where $\mu'(t) = \mu(t) + \bar{i}(t) - \bar{u}(t) - k(t)(\mu(t)-\bar{x}(t))$ (8.6)

which is the linear decision rule

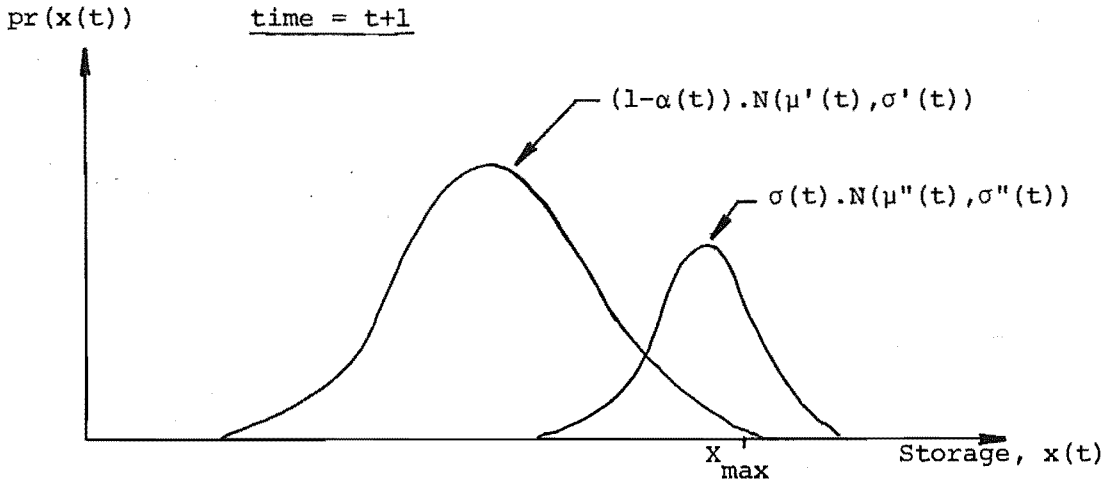
$$\sigma'^2(t) = \sigma_I^2(t) + (1-k(t))^2\,\sigma^2(t) \tag{8.7}$$

$$\mu''(t) = X_{max} + \bar{i}(t) - \bar{u}(t) - k(t)\,(X_{max}-\bar{x}(t)) \tag{8.8}$$

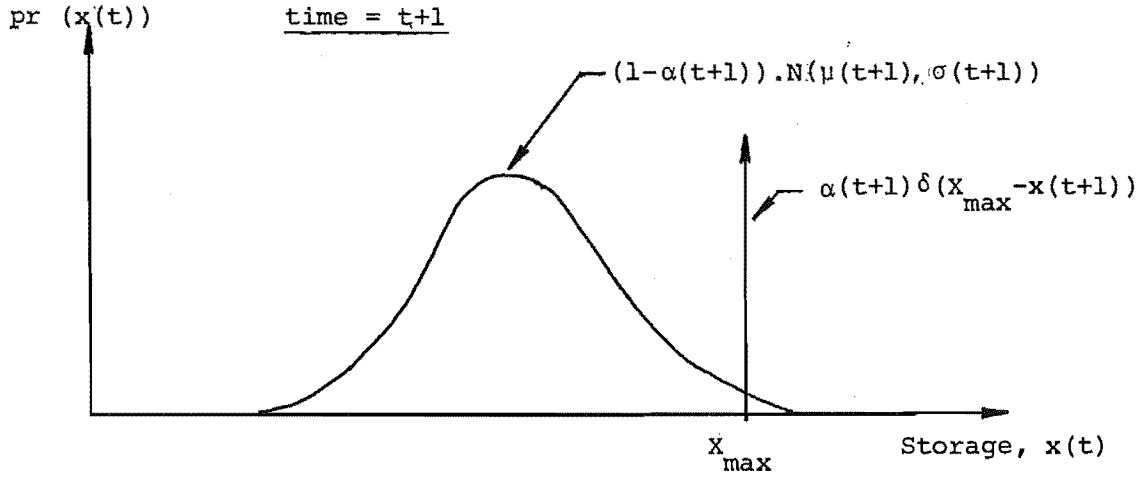
$$\sigma''^2(t) = \sigma_I^2(t) \tag{8.9}$$



(a) Storage distribution at time t, Gaussian + Impulse



(b) Storage distribution at time t+1 after adding Gaussian inflow and deducting releases from time t distributions



(c) Approximation to (b), new Gaussian + Impulse.

FIGURE 8.2 Storage probability density function evolution from t to t+1, with Gaussian + Impulse approximation.

and $\bar{i}(t)$ = mean inflow

$\sigma_I^2(t)$ = variance of inflows

$\bar{u}(t)$ = release for mean storage,

$\bar{x}(t)$ = mean storage,

U_{max} = maximum release

$k(t)$ = gain factor.

Determine $\alpha(t+1)$:

$$\alpha(t+1) = (1-\alpha(t))\frac{1}{\sigma'(t)\sqrt{2\pi}} \int_{X_{max}}^{\infty} e^{-(z-\mu'(t))^2/2\sigma'(t)^2} dZ$$

$$+ \alpha(t) \frac{1}{\sigma''(t)\sqrt{2\pi}} \int_{X_{max}}^{\infty} e^{-(Y-\mu''(t))^2/2\sigma''(t)^2} dY$$

$$= \frac{1}{2}\left[(1-\alpha(t))\ \text{erfc}\left(\frac{X_{max}-\mu'(t)}{\sigma'(t)\sqrt{2}}\right)\right.$$

$$\left. + \alpha(t)\ \text{erfc}\left(\frac{X_{max}-\mu''(t)}{\sigma''(t)\sqrt{2}}\right)\right] \qquad (8.10)$$

where the normal distribution is represented by an error function erf(x) and erfc(x)=1.-erf(x), (see Abramowitz & Stegun). Determine $\mu(t+1)$ for the two truncated Gaussians, i.e. for storages below X_{max}:

$$\mu(t+1) = \frac{1-\alpha(t)}{1-\alpha(t+1)} \int_{-\infty}^{X_{max}} \frac{Z}{\sigma'(t)\sqrt{2\pi}} e^{-(z-\mu'(t))^2/2\sigma'(t)^2} dZ$$

$$+ \frac{\sigma(t)}{1-\alpha(t+1)} \int_{-\infty}^{X_{max}} \frac{Y}{\sigma''\sqrt{2\pi}} e^{-(Y-\mu''(t))^2/2\sigma''(t)^2} dY \qquad (8.11)$$

Note that the factor $1/(1-\alpha(t+1))$ is necessary to form a valid probability density function of area 1.0 from the two Gaussian components, over the range of storages from $-\infty$ to X_{max}.

$$\therefore \quad \mu(t+1) = \frac{1-\alpha(t)}{1-\alpha(t+1)}\left[\frac{-\sigma'(t)}{\sqrt{2\pi}} e^{-\left(\frac{X_{max}-\mu'(t)}{2\sigma'(t)^2}\right)^2}\right.$$

$$\left. + \frac{\mu'(t)}{2}\left[1+\text{erf}\left(\frac{X_{max}-\mu'(t)}{\sigma'(t)\sqrt{2}}\right)\right]\right]$$

$$+\frac{\alpha(t)}{1-\alpha(t+1)}\left[-\frac{\sigma''(t)}{\sqrt{2\pi}}\,e^{\frac{-(X_{max}-\mu''(t))^2}{2\sigma''(t)^2}}+\frac{\mu''(t)}{2}\left[1+\mathrm{erf}\left(\frac{X_{max}-\mu''(t)}{\sigma''(t)\sqrt{2}}\right)\right]\right] \quad (8.12)$$

(See Appendix III for details of this integration).

Then find $\sigma^2(t+1)$ from the same two distributions using the formula

$$\sigma^2 = \int x^2\,p(x)\,dx - \mu^2$$

$$\sigma^2(t+1) = \frac{1-\alpha(t)}{1-\alpha(t+1)}\left[\int_{-\infty}^{X_{max}} \frac{Z^2}{\sigma'(t)\sqrt{2\pi}}\,e^{-(z-\mu'(t))^2/2\sigma'(t)^2}\,dZ\right]$$

$$+\frac{\sigma(t)}{1-\alpha(t+1)}\left[\int_{-\infty}^{X_{max}} \frac{Y^2}{\sigma''(t)\sqrt{2\pi}}\,e^{-(Y-\mu''(t))^2/2\sigma''(t)^2}\,dY\right] - \mu^2(t+1)$$

$$=\frac{1-\alpha(t)}{1-\alpha(t+1)}\left[\frac{\sigma'(t)^2}{2}-\frac{\sigma'(t)(X_{max}+\mu'(t))}{\sqrt{2\pi}}\,e^{-(X_{max}-\mu'(t))^2/2\sigma'(t)^2}\right.$$

$$\left.+\frac{\mu'(t)^2+\sigma'(t)^2}{2}\,\mathrm{erf}\left(\frac{X_{max}-\mu'(t)}{\sigma'(t)\sqrt{2}}\right)+\frac{\mu'(t)^2}{2}\right]$$

$$+\frac{\sigma(t)}{1-\sigma(t+1)}\left[\frac{\sigma''(t)^2}{2}-\frac{\sigma''(t)(X_{max}+\mu''(t))}{\sqrt{2\pi}}\,e^{\frac{-(X_{max}-\mu''(t))^2}{2\sigma''(t)^2}}\right.$$

$$\left.+\frac{\mu''(t)^2+\sigma''(t)^2}{2}\,\mathrm{erf}\left(\frac{X_{max}-\mu''(t)}{\sigma''(t)\sqrt{2}}\right)+\frac{\mu''(t)^2}{2}\right]-\mu^2(t+1) \quad (8.13)$$

(Details of this integration are given in Appendix III).

Constraints on minimum values of storage and release can not be handled in the same way as the simpler Gaussian model of Chapter 7. The probability of being below some level $\mu(t)-a\sigma(t)$ now depends on $\alpha(t)$. Constraints must now be formulated in terms of probabilities of violation allowable, instead, i.e. the following quantities are added to the cost function:

States:

$$W_x\left[(1-\alpha(t)).\frac{1}{2}\left(1+\mathrm{erf}\left(\frac{X_{min}-\mu(t)}{\sqrt{2}\,\sigma(t)}\right)\right)-P_x\right]^2 \quad (8.14)$$

$$\text{if } (1-\alpha(t)).\frac{1}{2}.\left(1+\mathrm{erf}\left(\frac{X_{min}-\mu(t)}{\sqrt{2}\,\sigma(t)}\right)\right) > P_x$$

where P_x = acceptable probability of violating state constraint.

Controls:

Control constraint is violated if

$$U_{min} \gtrsim \bar{u}(t) + k(t)(x(t)-\bar{x}(t)) \tag{8.15}$$

$$\Rightarrow x(t) \lesssim \frac{(U_{min}-\bar{u}(t))}{k(t)} + \bar{x}(t) \tag{8.16}$$

for the constraint to be violated if $k(t) > 0$ or

$$x(t) \gtrsim \frac{(U_{min}-\bar{u}(t))}{k(t)} + \bar{x}(t) \quad \text{if } k(t) < 0. \tag{8.17}$$

The penalty is now:

$$W_u \left((1-\alpha(t)).\frac{1}{2}.\left[1+\text{Sgn}(k)\,\text{erf}\left(\frac{\frac{U_{min}-\bar{u}(t)}{k(t)} - \mu(t)}{\sqrt{2}\,\sigma(t)}\right)\right] - P_u \right)^2 \tag{8.18}$$

$$\text{if } (1-\alpha(t)).\frac{1}{2}.\left[1+\text{Sgn}(k)\,\text{erf}\left(\frac{\frac{U_{min}-\bar{u}(t)}{k(t)} - \mu(t)}{\sqrt{2}\,\sigma(t)}\right)\right] > P_u$$

where $\text{Sgn}(k) = 1$ if $k(t) \gtrsim 0$

or $\quad -1$ if $k(t) < 0$.

In addition to the penalties of eqns (8.14) and (8.18), separate penalties are applied to $\mu(t)$ and $\bar{u}(t)$. If the probability of violating a constraint is approximately 1.0, then that penalty will not affect the gradient, i.e. the mean value of that quantity could violate the constraint to a greater extent without increasing the probability of constraint violation. To overcome this difficulty the following two extra penalty terms are added to the cost function:

$$W_x \left((1-\alpha(t)) \left(\frac{X_{min}-\mu(t)}{X_{max}} \right) \right)^2 \tag{8.19}$$

$$\text{if } \mu(t) < X_{min}$$

$$\text{and} \quad \left(W_u \left((1-\alpha(t)) \left(\frac{\frac{U_{min}-\bar{u}(t)}{k(t)} - \mu(t)}{X_{max}} \right) \right) \right)^2 \tag{8.20}$$

$$\text{if } \bar{u}(t) < U_{min}.$$

The upper constraint on storage is handled by the "impulse" part of the model. Releases above the maximum are permitted but do not reduce thermal fuel costs any more than the maximum release itself, as for the Gaussian model (eqn (7.19) applies).

Fuel costs were also calculated in the same way for the Gaussian component, scaled by $1 - \alpha(t)$. The cost applicable to maximum storages is calculated separately, the single value for that point simply being multiplied by its probability $\alpha(t)$. Otherwise, the only modification required was a more accurate formula for the error function than that of eqn (7.22) which gave $|\varepsilon(x)| \leq 5 \times 10^{-4}$ i.e.:

$$\text{erf}(x) = 1. - \frac{1}{\left[1+a_1x+a_2x^2+..+a_6x^6\right]^{16}} + \varepsilon(x) \qquad (8.21)$$

where $|\varepsilon(x)| \leq 3 \times 10^{-7}$

$a_1 = 07052\ 30784$ $\quad a_2 = 04228\ 20123$

$a_3 = 00927\ 05272$ $\quad a_4 = 00015\ 20143$

$a_5 = 00027\ 65672$ $\quad a_6 = 00004\ 30638$

This change was made as test runs showed a significant difference in the solutions obtained from the two formulas.

8.2.3 Finite Differences Gradients

Explicit gradient equations and their evaluation were avoided by using a finite differences method. This requires the perturbation of each control variable in turn, and the calculation of the resulting function value to give gradients. So to find the function value and gradient at a single point requires a total of (26xN)+1 function evaluations, for N cycles, i.e. the first to find the function value at that point, then 26 for each cycle to find the gradients. Obviously this is a very inefficient process, so far as computer time is concerned. It permits more rapid development of an algorithm especially when many modifications are made. Now there are no gradient and costate calculations to alter. The hill climbing algorithm, CGRADS was used without modification with these gradients.

At first the Hamiltonian method used for the deterministic problem was applied to the stochastic optimisation. Derivatives with respect to $\sigma^2(t)$, $\mu(t)$ and $\alpha(t)$ are required for costates ($\partial H/\partial$state), and with respect to $u(t)$ and $k(t)$ for gradients, ($\partial H/\partial$control). Performing the necessary algebra and programming accurately was a formidable task. The great increase in computer time required seemed a small price to pay for the simplifications given by the finite differences method.

Choice of a perturbation size was guided by the principle that it should be less than the desired accuracy for the controls.

8.3 RESULTS FROM THE NEW MODEL

8.3.1 Preliminaries

The results will be compared with stochastic D.P. solutions, as usual. It will be shown that adding an impulse to the Gaussian of the previous linear feedback model gives significantly better results. Firstly the different basis for assessing state constraint violations is examined.

When representing storage probability distributions simply by a Gaussian, the penalty for state constraints is applied if:

$$\mu - 1.5\ \sigma_x < X_{min}$$

i.e. if $Pr(x < X_{min}) > 0.0668$

The allowable violation for convergence is:

$$\mu - 1.5\ \sigma_x > X_{min} - 0.05\ X_{max}$$

For the Gaussian plus impulse model, a penalty is applied when:

$$Pr(x < X_{min}) > 0.0670$$

and for convergence, the requirement is that

$$Pr(x < X_{min}) < 0.0800$$

For the run of Figure 7.10 (3 cycles, initial storage = final storage = 2250), the maximum constraint violation on the first cycle gave

$$\mu - 1.46\ \sigma_x = X_{min}$$

$$\Rightarrow Pr(x < X_{min}) = 0.0720$$

The equivalent run of the Gaussian plus impulse method gave a maximum probability of violation of 0.0689. So the minimum constraint is not violated to quite such an extent.

8.3.2 Convergence

Convergence is improved by the addition of the impulse. From Table 8.1, the changes in initial (t=1) release decisions are insignificant for a 25% reduction in final time target storage. Even for one cycle (one year only) the effect is only 0.001%. Processing time increases with the square of the problem size (number of cycles), due to the finite differences gradients calculations. With analytic gradients, process time would go up in proportion to the number of cycles only.

The results to be analysed will be taken from two cycle runs however. The reason is apparent from Table 8.2, showing mean storages for the first cycle of one, two and three cycle runs. Two and three cycle solutions differ by at most 0.05% (for t=11), indicating that the first cycle of the two cycle run is on the steady state trajectory, i.e. the effect of the arbitrary final time storage is negligible. The one cycle run differs from the three cycle case significantly at t=4 (0.05%).

8.3.3 The Optimal Strategy

The first two comparisons with S.D.P. are of the most important result - the initial release decision, which is the only part of the solution to be actually implemented. Releases for t=1 for various storage levels are compared in Figure 8.3 (cf. Fig. 7.8 for the Gaussian version). The desired improvement in the approximation at larger releases is evident, compared with the Gaussian only model. The latter effectively had extra storage and so retained more water for use later, at peak demand times. The releases for the Gaussian plus impulse are a little larger than the S.D.P. values for large storages where the S.D.P. curve is rising steeply.

The optimisation was begun at t=6 in Figure 8.4. All data have been rotated back in time five steps for these runs. These results are what would be actually used at t=6 when implementing the algorithm as an open loop feedback controller. This is a high demand, low storage time, with releases at their maximum value when the reservoir is above about 60% full - quite a different pattern to that for t=1 and so a good test. The Gaussian plus impulse results follow the S.D.P. solution even better than at t=1.

TABLE 8.1 Releases for t=1 for Various Runs for Gaussian Plus Impulse Linear Feedback Model, to Show Convergence. EPSU (control variable convergence tolerance) = 0.031. Perturbation for Gradient Calculations = 0.02.

Number of Cycles	Initial Storage = 2250 Mean Final Storage: 2250	1690	Computer Process Time For Initial Storage = 2250
1	838.71	838.70	1 hr 45 min
2	838.71	838.83	6 hrs 13 min
3	838.72	838.71	14 hrs 46 min

TABLE 8.2 Mean Storages, for First Cycle in Each Case, Showing Convergence of Gaussian Plus Impulse Model

Time	1 Cycle	2 Cycles	3 Cycles
1	2250.0	2250.0	2250.0
2	2435.8	2435.8	2435.8
3	2424.2	2423.2	2423.4
4	2146.1	2143.9	2143.9
5	1675.2	1673.5	1673.8
6	1258.4	1259.2	1259.1
7	908.7	914.3	914.0
8	896.6	911.4	910.8
9	956.8	998.8	999.5
10	1184.9	1287.2	1287.9
11	1568.8	1717.0	1717.9
12	1918.9	2099.7	2100.7
13	2121.9	2308.4	2308.2
14	2155.6	2340.0	2339.2

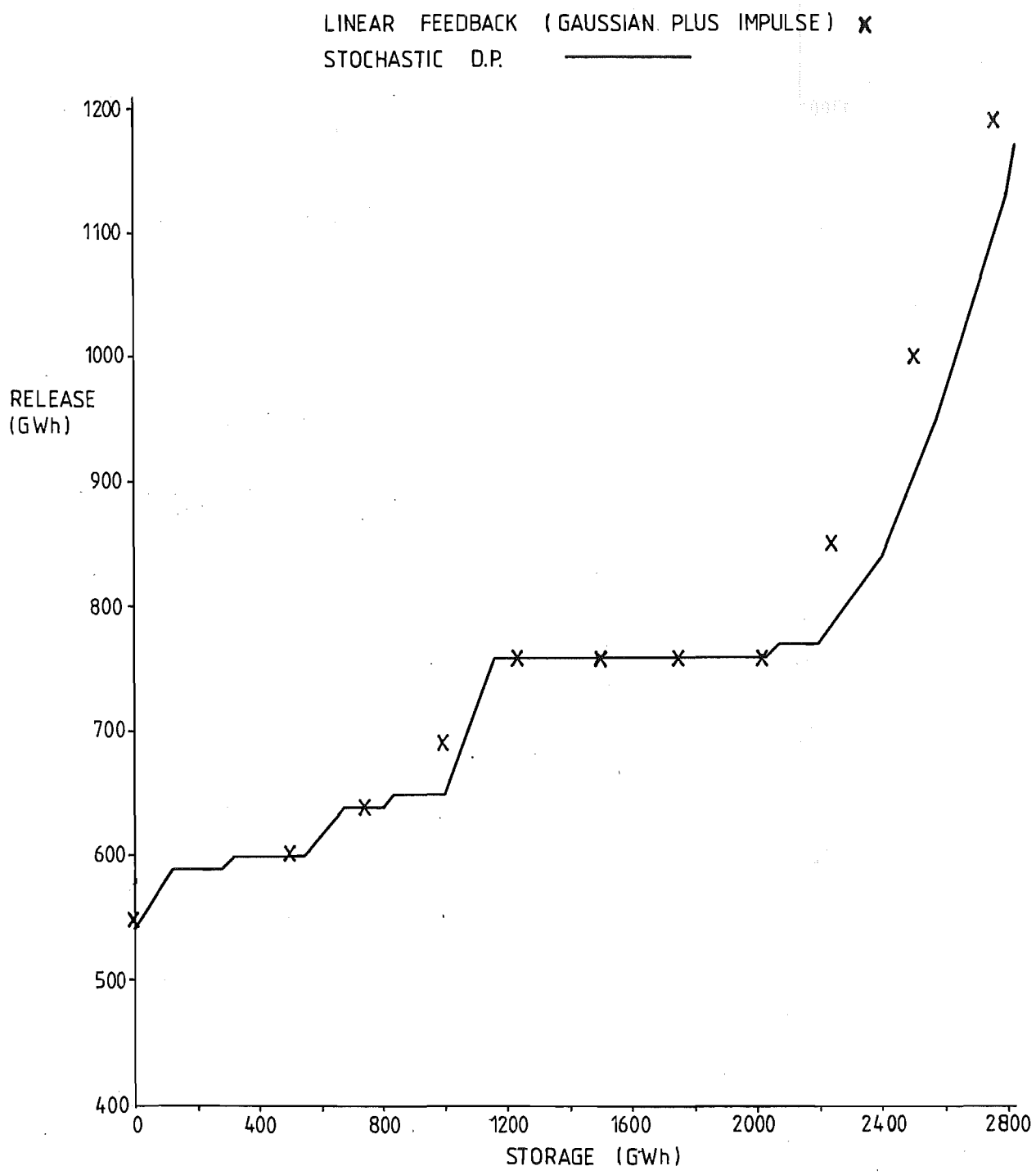


FIGURE 8.3 Releases for t=1 (first time interval) for Gaussian plus impulse and S.D.P.

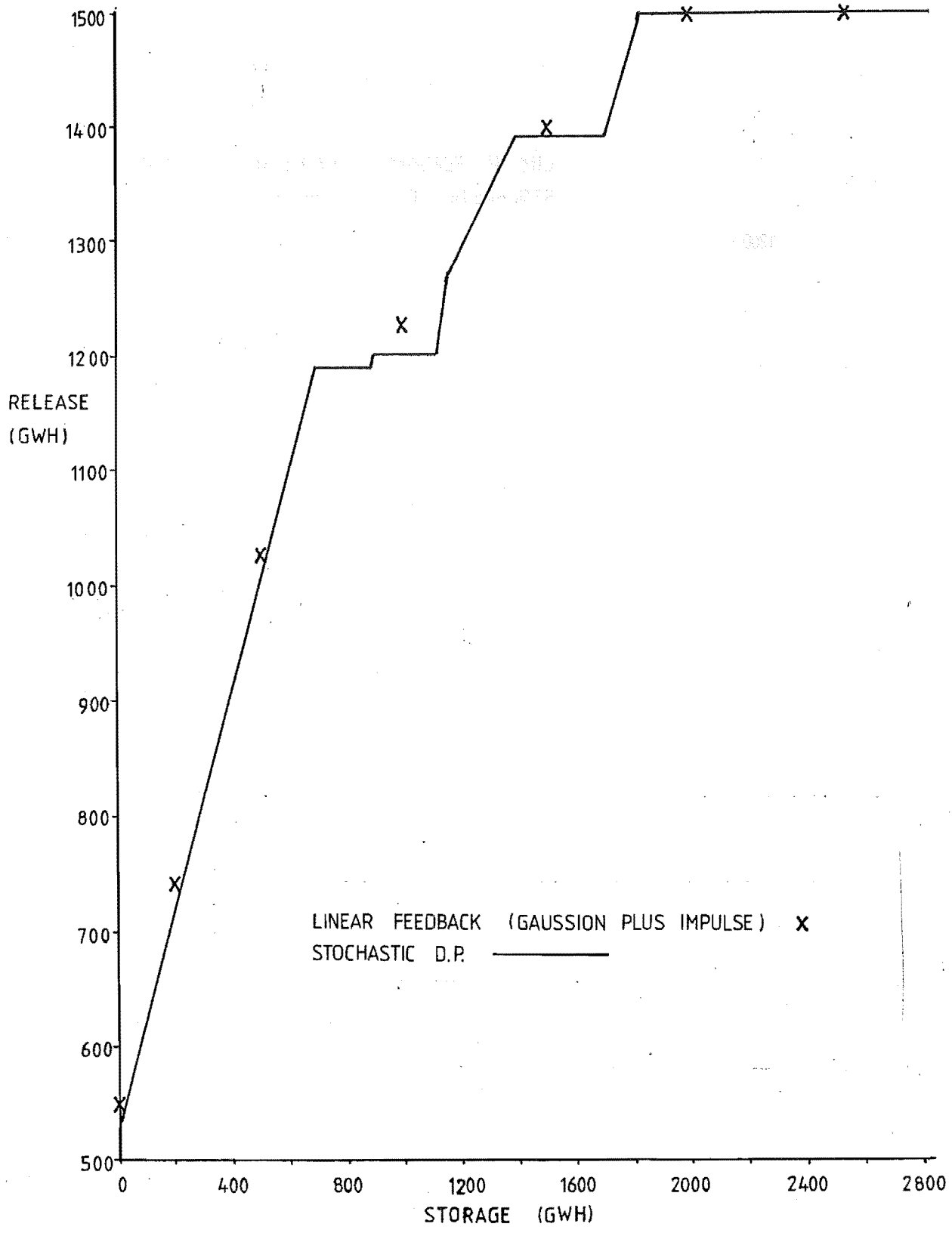


FIGURE 8.4 Releases for Gaussian plus impulse and S.D.P., when beginning optimisation at t=6 (i.e. moving data round 5 steps).

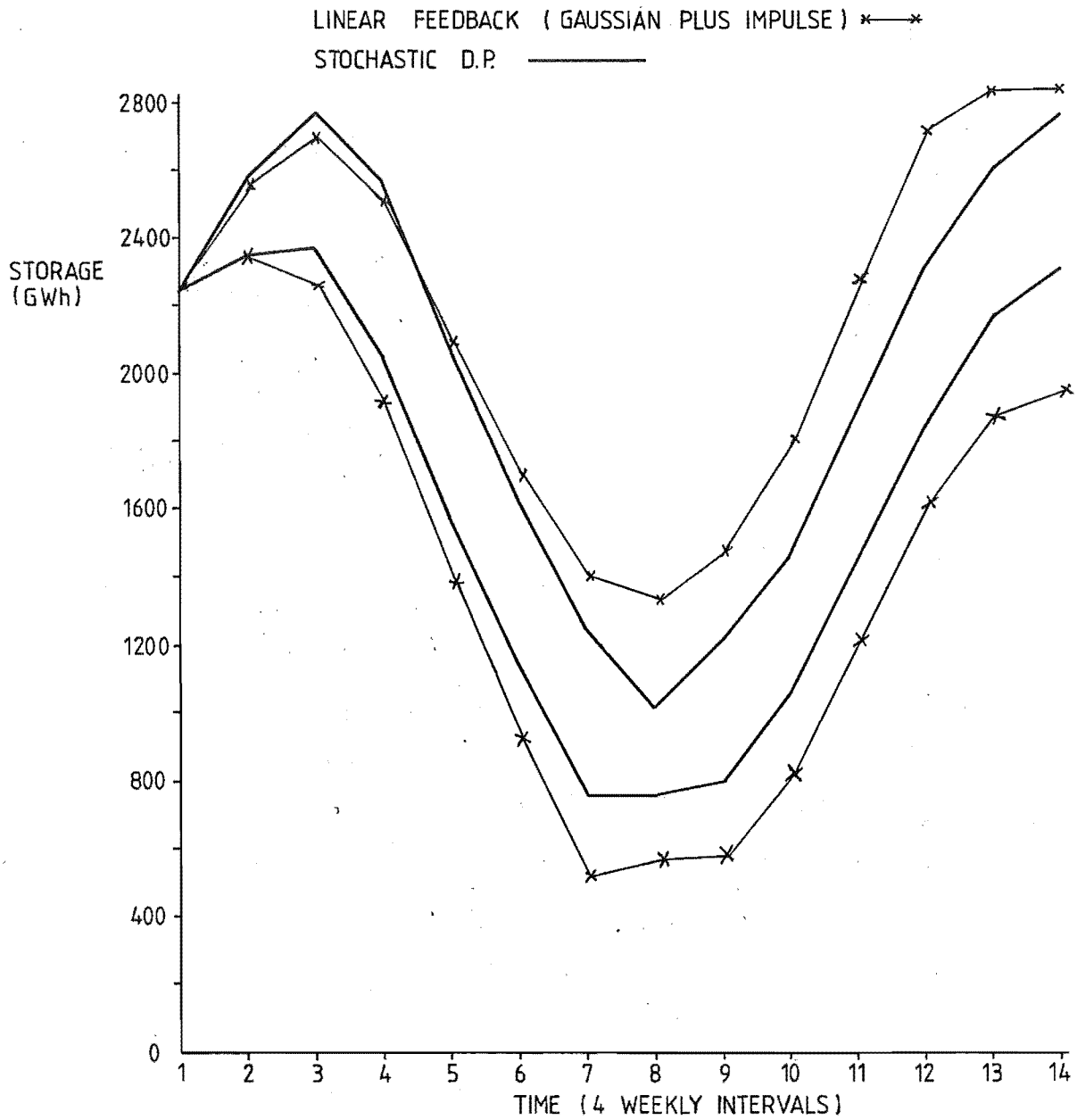


FIGURE 8.5 Mean ± σ/3 Flows storage trajectories for S.D.P. and Gaussian plus impulse linear feedback model.

TABLE 8.3 Storages for Mean Inflows for the Three Stochastic Methods. Note: SDP Discretised to 10 GWh Steps.

Time	Gaussian Plus Impulse	Gaussian Only	Stochastic Dynamic Program
1	2250	2250	2250
2	2454	2456	2460
3	2486	2541	2550
4	2208	2247	2270
5	1734	1775	1790
6	1314	1342	1370
7	962	983	950
8	952	978	860
9	1033	1147	1030
10	1318	1483	1370
11	1754	1969	1770
12	2165	2423	2220
13	2440	2726	2480
14	2500	2828	2630

When analysing the earlier Gaussian only method, the mean storage trajectory was compared with the S.D.P. result in Figure 7.10. This time such a comparison is not worthwhile as the two are very close over most of their length. Instead the close agreement is shown in Table 8.3. Note that "mean storages" are not the same as "storages with mean inflows" for the Gaussian plus impulse method. The largest deviation from the S.D.P. trajectory, as percentages of maximum storage, are 4.6% at t=14 and 3.2% at t=8 (130 GWh and 92 GWh respectively.).

The improvement obtained by adding the impulse to the Gaussian method is obvious from Table 8.3. The trajectory is no longer consistently above the S.D.P. for the last few time intervals. The reverse occurs, but only to a small extent with new version. Agreement with the S.D.P. at t=8, when storages are low, will be determined largely by the choice of allowable minimum storage constraint violation.

Mean $\pm \sigma/3$ inflow storage trajectories are shown in Figure 8.4 (Fig. 7.11 for Gaussian only model). The pattern is changed only a little, the spread being reduced slightly for the second half of the year. Final time

spread in storages is virtually the same for the two models. Gaussian plus impulse storages for high inflows are on the maximum level for the last time interval only, as against the last two for the Gaussian model.

8.3.4 Simulation Check

The next assessment of results is by comparison of the storage probability distributions given by the Gaussian plus impulse with those given by a more accurate simulation. This simulation uses the same control strategy but discretises storage at each time into 285 steps of 10 GWh each, the same as the D.P. methods. Starting at the known initial storage this procedure models the evolution of the probability distribution exactly, subject only to the errors in discretising to 10 GWh increments. Table 8.4 compares α (the probability of being at maximum storage), mean storage and standard deviation for the simulation and the Gaussian plus impulse approximation. The two correspond remarkably well - the greatest disagreements are:

(i) α, .027 for t=14,

(ii) μ, 79 or 5.8% for t=10,

(iii) σ, 37 or 5.8% for t=9.

These results indicate that the very simple Gaussian sum used gives an adequate representation of the distribution, despite the distortions caused by truncating the Gaussian at maximum storage every time interval.

8.3.5 Fuel Costs

A final means of evaluating the extent of sub-optimality of our solution method is by the fuel costs of Table 8.5. Table 8.5(b) can be compared with 7.7(b). Agreement in all cases is good for mean storages, but the Gaussian plus impulse costs are much lower for high inflows, than the simple Gaussian model's result, i.e. the new method reduces waste of higher inflows. For low inflows both linear feedback methods give similar results to the S.D.P., whereas the deterministic D.P. does not.

8.4 CONCLUSIONS

The finite differences gradient technique has been shown to work effectively with a standard algorithm designed for use with analytic gradients. It is a good means of simplifying development when gradients are difficult to calculate, and the task of interest is to test a solution approach, rather than the development of the most efficient algorithm possible.

TABLE 8.4 Gaussian Plus Impulse Method Results for 1st of 2 Cycles and Comparison of Storage Probability Distribution Evolution with that of a more Accurate Simulation Using the Same Control Strategy

	Gaussian Plus Impulse Optimisation					Exact Simulation of Control Strategy		
	Mean Release	Gain	α	Mean Storage	Standard Deviation	α	Mean Storage	Standard Deviatic
1	838.7	0	0	2250	0	0	2250	0
2	1021.4	.2068	.116	2436	274	.116	2432	274
3	1274.3	.0967	.199	2423	341	.203	2423	345
4	1429.6	.0798	$.538\times10^{-1}$	2144	407	$.518\times10^{-1}$	2155	416
5	1346.0	.100	$.333\times10^{-2}$	1674	452	$.108\times10^{-2}$	1695	460
6	1285.3	.1194	$.378\times10^{-3}$	1259	472	$.555\times10^{-4}$	1292	474
7	1137.7	.1484	$.543\times10^{-4}$	914	495	$.128\times10^{-4}$	963	482
8	1160.2	.1659	$.838\times10^{-3}$	911	613	$.929\times10^{-3}$	979	580
9	1038.3	.1696	$.228\times10^{-2}$	999	640	$.266\times10^{-2}$	1077	603
10	879.1	.1829	$.854\times10^{-2}$	1287	627	$.101\times10^{-1}$	1366	598
11	845.4	.1985	$.445\times10^{-1}$	1717	590	$.509\times10^{-1}$	1793	567
12	950.9	.2300	.124	2100	495	.130	2167	483
13	1023.9	.2658	.233	2308	462	.250	2356	452
14			.216	2340	428	.243	2384	426

TABLE 8.5 Comparison of Fuel Costs from Stochastic D.P. and Gaussian Plus Impulse (2 cycles, initial and final storages = 2250 GWh)

(a)

Inflows	Stochastic D.P.		Gaussian + Impulse	
	Cost ($\$x10^6$)	Final Storage (GWh)	Cost ($\$x10^6$)	Final Storage (GWh)
$\mu - \sigma/3$	66.6	2310	59.5	1947
μ	48.5	2630	46.6	2500
$\mu + \sigma/3$	30.8	2770	34.5	2850

(b)

Inflows	Total Annual Fuel Cost + Stochastic D.P. Final Time Storage Penalty ($\$x10^6$)	
	Stochastic D.P.	Gaussian + Impulse
$\mu - \sigma/3$	75.2	76.9
μ	51.5	51.6
$\mu + \sigma/3$	31.9	34.5

Adding the impulse of probability at maximum storage has overcome the difficulties at high storages evident in the Gaussian linear feedback model of Chapter 7. Agreement with S.D.P. results is now good for high storages, not just for lower storages, as with the simpler model. Only the initial release decision would be used in practice, as the algorithm would be used as an open loop feedback controller. Hence the validity of this decision is crucial to the usefulness of any method. Two different times during the year for which the S.D.P. gives quite different release versus storage patterns have been tested to examine the initial release decisions from the new algorithm. In both cases our sub-optimal method reproduced the true S.D.P. results very satisfactorily.

Computer time requirements might seem high at over six hours for a two cycle problem, which gives a 54 dimensional static optimisation problem. If explicit gradient calculations were used, the subroutine calculating gradient and function value might require twice as much computation as the one used to calculate only a single function value. Hence total processing time would drop to about 15 minutes, or about 1/26th of the total figure quoted above.

Some suggestions for extension of this algorithm are left for the next chapter, the conclusions to this thesis. This is the area that has plenty of scope for development of new ideas, to obtain a useful multi-reservoir scheduling method overcoming the S.D.P. dimensionality problem. The deterministic solutions on the other hand are less realistic but obtainable by established techniques, so do not feature in the suggestions for further research. The objective, ideally, is to incorporate the stochastic features of the Gaussian plus impulse model into the large, detailed eight reservoir deterministic model.

CHAPTER 9

CONCLUSIONS

This chapter answers three questions:

(i) What has this thesis achieved?

(ii) How could this work be used as an aid to power system operation?

(iii) What further development of these ideas would be useful?

9.1 ACHIEVEMENTS

The first conclusion drawn during the research for this thesis was that optimisation by decomposition is not always a sufficiently reliable technique, or as advantageous as was first thought. Strict convexity of the problem is essential, and for good convergence properties the boundaries of the feasible region, or the cost function, should not be nearly straight, i.e. the more curvature the better.

The first really useful achievement was the development of a reliable conjugate gradients method. It is designed not to hang up on minor irregularities in the function being minimised, and to give accurate solutions. It performed well on a number of problems, for some of which analytic gradient calculations were possible, but others used a finite differences method.

A model for the deterministic optimal scheduling of hydro thermal power systems was proposed and successfully applied to a 312 dimensional, 8 reservoir problem. Conventional non-linear programming techniques worked well on this large problem, and it was possible to incorporate all factors considered important for the present manual scheduling methods.

The first conclusion regarding the stochastic problem is that stochastic and deterministic (open loop feedback) solutions are significantly different, (see Figures 7.6, 7.7 and 7.8). The consequences of these differences on the usefulness of the algorithms is discussed in the next section. Clearly efforts to develop a stochastic method are justified.

Hydro reservoir optimisation is peculiar in that lower limits on storage and water release must be enforced, but upper constraints can be handled by spilling. This is realistic in that in practice excess storage or release corresponds to water wasted - lost from the system without any direct benefits. Modelling of this feature sets the problem apart from other similar efforts in the field of reservoir management. The "Gaussian plus impulse" probability distribution for water storages models this behaviour. The concept was suggested by the Gaussian sums principle, but our scheme gives an economical strategy for handling a non-Gaussian variable, for this particular application.

Coupled with linear feedback, the Gaussian plus impulse probability distribution led to the successful solution of a one reservoir test system. Ideally the results should imitate the true solution obtained by stochastic dynamic programming. Agreement of the releases for the first time interval for various storage levels, was especially good (see Fig's 8.3 and 8.4). This result is crucial, as it is the only decision actually used, in practice. Evolution of the storage distribution for the new method also follows the S.D.P. results well.

While S.D.P. gives the true solution, the curse of dimensionality prevents its use for multi-reservoir problems. The new method proposed does not suffer from this limitation.

9.2 IMPLEMENTATION

The general principles set out in Chapter 3 could be applied to most power systems, but perhaps using some other approximation for thermal fuel costs. The New Zealand model requires some further work before it could be utilised. Three load segments were used to represent load diversity, but checks with perhaps 5 and 7 segments would be needed to determine the required number. Weekly time intervals, rather than 4 weeks, are probably required, especially if the algorithm is to be used as a reservoir management aid. It would of course be used in the open loop feedback mode in this case, resolving for a whole year's operation after each week.

Detail adjustments to the model include use of actual station maintenance schedules, with up to date data for forced outage probabilities, fuel costs, and inflows. Demand predictions, not historical values should be

used. It could then become a very useful tool for investigation of system performance under various conditions of demand or inflow, or for examining system development options. Other uses include incremental costs of supply, and fuel cost estimation.

Application to actual reservoir scheduling is more difficult. Comparison of one reservoir stochastic and deterministic D.P. solutions showed the considerable differences in strategy. For example, consider the situation if low water storage was to occur in mid-winter. On the basis of mean inflows, the deterministric algorithm might suggest a stragegy that would empty all reservoirs. It would not leave extra water in storage to cope with less than mean inflows. The system control staff would then probably reject this strategy and use more thermal generation. This would reduce the probability of running out of water, and of possibly suffering a shortfall in energy supplies. For scheduling purposes, the deterministic strategy is therefore more likely to be used to help examine the effects of various inflow levels. It is unlikely to be more than just another factor in the decision making process.

Hence further development to achieve a practicable stochastic scheduling method is justified.

9.3 FURTHER RESEARCH

The effectiveness of the Gaussian plus impulse, linear feedback method on a one reservoir problem has been shown clearly. The requirement for scheduling to be by a stochastic method has been discussed. Before developing this method further, some modifications might be investigated. A piecewise linear feedback rule, with different gains applying above and below the mean storage could more accurately imitate the S.D.P. It might help when the mean release equals, or is near to, either its maximum value or the demand. In this situation releases might exhibit a saturation effect i.e. releases increase rapidly with storage until the mean storage is reached, but the release is constant above the mean storage. Dynamic programming release against storage plots (Figure 7.9) exhibited behaviour of that type.

The major step to a two or three reservoir model would be very interesting. For the New Zealand model, an examination of correlation of inflows between the eight reservoirs would indicate the number needed. If this step could be taken, and reliable solutions obtained, then a really useful sched-

uling method would be possible. Such an algorithm might provide decisions that system operators could have some faith in. Only the choice of acceptable probabilities of running out of water would still rely on human judgement.

APPENDIX I

LISTING OF CGRADS CONJUGATE GRADIENTS COMPUTER PROGRAM - as used on VAX 11/780 computer, FORTRAN IV programming

```
100   C*********************************************************************
200   C
300   C              SUBROUTINE CGRADS2
400   C
500   C  FINDS LOCAL MINIMUM OF A FUNCTION USING GENERALISED CONJUGATE
600   C  GRADIENTS METHOD WITH BEALE RESTARTS.   SEPTEMBER 1981
700   C
800   C  LINEAR SEARCH ROUTINE CONTROLLED BY GRADIENTS ONLY
900   C
1000  C  MAGNITUDES OF SCALAR PARAMETERS USED IN CALCULATION OF
1100  C   SEARCH DIRECTIONS ARE LIMITED TO AVOID VERY LARGE COMPONENTS
1200  C   IN SEARCH DIRECTION VECTORS.
1300  C
1400  C  4 TYPES OF SEARCH DIRECTION POSSIBLE:
1500  C    II=1  RS=1  STEEPEST DESCENT
1600  C    II-1  RS=0  STEEPEST DESCENT STARTED, CONJUGATE DIRECTION
1700  C    II=0  RS=1  BEALE RESTART
1800  C    II=0  RS=0  BEALE RESTARTED, CONJUGATE DIRECTION
1900  C
2000  C
2100  C
2200  C*********************************************************************
2300  C
2400        SUBROUTINE CGRADS (FUNCT,N,X,F,G,EPSX,EPSG,LIMIT,IER,SD,NFE,GL,
2500       *GDRS,SDRS,XLS,SZR,RSC,KOUNT,SFACT)
2600        IMPLICIT REAL*8 (A-H,O-Z)
2700  C
2800  C  MATRIX ICG RECORDS TYPE OF SEARCH USED EACH ITERATION AND REASON
2900  C  FOR EACH ITERATION TERMINATING.
3000  C     ICG(KOUNT,1) = NUMBER OF FUNCTION EVALUATIONS AT END OF ITERATION
3100  C                          NUMBER KOUNT.
3200  C     ICG(KOUNT,2) = II VALUE
3300  C     ICG(KOUNT,3) = RS VALUE
3400  C     ICG(KOUNT,4) = POINT AT WHICH LINEAR SEARCH FOR THIS ITERATION
3500  C                            ENDED
3600        COMMON/BLK4/ICG(1000,4),FCG(2,1001)
3700  C
3800  C  DIMENSIONED DUMMY VARIABLES
3900        REAL*8 X(1),G(1),SD(1),GL(1),GDRS(1),SDRS(1)
4000        REAL*8 LSZ,XOLD(312)
4100        INTEGER RS
4200  C
4300  C  RESET COUNTERS ETC
4400        NFE=0
4500        KOUNT=0
4600        II=1
4700        RS=1
4800        IER=0
4900        DO 5 I=1,N
5000        XOLD(I)=X(I)
5100     5  CONTINUE
5200  C
5300  C  COMPUTE FUNCTION VALUE AND GRADIENT FOR INITIAL ARGUMENT
```

```
5400          CALL FUNCT(X,F,G,NFE)
5500          FCG(1,1)=F
5600          FCG(2,1)=0.
5700          FOLD=F
5800    C  INCREMENT ITERATION COUNTER
5900       10 KOUNT=KOUNT+1
6000    C  COMPUTE SQUARE OF GRADIENT
6100          GNRM=0.
6200          DO 20 J=1,N
6300          GNRM=GNRM+G(J)*G(J)
6400       20 CONTINUE
6500    C  IS THIS A RESTART STEP?
6600          IF(RS.NE.1)GO TO 100
6700    C  IS THIS A STEEPEST DESCENT STEP?
6800          IF(II=1)70,40,40
6900       30 II=1
7000          RS=1
7100    C
7200    C  STEEPEST DESCENT SEARCH DIRECTION
7300       40 DY=0
7400          DO=50 J=1,N
7500          SD(J)=-G(J)
7600          SD(J+N)=X(J)
7700          DY=DY+SD(J)*G(J)
7800       50 CONTINUE
7900          GO TO 200
8000    C
8100    C  BEALE RESTART STEP
8200       60 II=0
8300          RS=1
8400       70 BETA1=0
8500          BETA2=0
8600          DO 80 J=1,N
8700          GDRS(J)=G(J)-GL(J)
8800          SDRS(J)=SD(J)
8900          BETA1=BETA1+G(J)*GDRS(J)
9000          BETA2=BETA2+SD(J)*GDRS(J)
9100       80 CONTINUE
9200          IF(BETA1*BETA2.EQ.0.)GO TO 30
9300          BETA=BETA1/BETA2
9400          IF(ABS(BETA).GT.1000.)GO TO 30
9500       CALCULATE SEARCH DIRECTION
9600          DY=0
9700          DO 90 J=1,N
9800          SD(J)=G(J)+BETA*SD(J)
9900          SD(J+N)=X(J)
10000         DY=DY+SD(J)*G(J)
10100      90 CONTINUE
10200   C  IS SEARCH DIRECTION SUFFICIENTLY DOWNHILL?
10300   C  GO TO STEEPEST DESCENT IF NOT
10400         IF(-1.2*GNRM.GT.DY)GO TO 30
10500         IF(DY.GT.-0.8*GNRM)GO TO 30
10600         GO TO 200
10700   C
10800   C  NON RESTART SEARCH DIRECTION CALCULATIONS.
10900   C
11000   C  IS A RESTART NEEDED?
11100      100 GT=0.
11200          DO 110 J=1,N
```

```
11300          GT=GT+GL(J)*(J)
11400        110 CONTINUE
11500          IF(GT.GE.RSC*GNRM)GO TO 60
11600    C  CALCULATE BETA
11700          BETA1=0.
11800          BETA2=0.
11900          DO 120 J=1,N
12000          GD=G(J)-GL(J)
12100          BETA1=BETA1+G(J)*GD
12200          BETA2=BETA2+SD(J)*GD
12300        120 CONTINUE
12400          IF(BETA1*BETA2.EQ.0.)GO TO 30
12500          BETA=BETA1/BETA2
12600          IF(ABS(BETA).GT.1000.)GO TO 30
12700          DY=0.
12800    C  WAS THIS SET OF CONJUGATE DIRECTIONS BEGUN BY A STEEPEST DESCENT
12900    C  STEP OR BY A BEALE RESTART?
13000          IF(II.EQ.1)GO TO 170
13100    C
13200    C  CALCULATE GAMA FOR BEALE RESTARTED SERIES
13300          GAMA1=0.
13400          GAMA2=0.
13500          DO 130 J=1,N
13600          GAMA1=GAMA1+G(J)*GDRS(J)
13700          GAMA2=GAMA2+SDRS(J)*GDRS(J)
13800        130 CONTINUE
13900          IF(GAMA1*GAMA2.EQ.0.)GO TO 30
14000          GAMA=GAMA1/GAMA2
14100          IF(ABS(GAMA).GT.1.E4)GO TO 30
14200          DO 140 J=1,N
14300          SD(J)=-G(J)+BETA*SD(J)+GAMA*SDRS(J)
14400          SD(J+N)=X(J)
14500          DY=DY+SD(J)*G(J)
14600        140 CONTINUE
14700    C  IS SEARCH DIRECTION SUFFICIENTLY DOWNHILL?
14800    C  RESTART SETTING GAMA=0 IF NOT
14900          IF(-1.2*GNRM.GT.DY)GO TO 150
15000          IF(DY.LT.-0.8*GNRM)GO TO 200
15100    C  RESTART
15200        150 II=0
15300          RS=1
15400          DY=0.
15500          DO 160 J=1,N
15600          SD(J)=SD(J)-GAMA*SDRS(J)
15700          GDRS(J)=G(J)-GL(J)
15800          SDRS(J)=(SD(J)+G(J))/BETA
15900          DY=DY+SD(J)*G(J)
16000        160 CONTINUE
16100    C  IS THIS RESTART DIRECTION SATISFACTORY?
16200    C  GO TO STEEPEST DESCENT IF NOT
16300          IF(-1.2*GNRM.GT.DY)GO TO 30
16400          IF(DY.GT.-0.8*GNRM)GO TO 30
16500          GO TO 200
16600    C
```

```
16700   C  CALCULATE SEARCH DIRECTION - STEEPEST DESCENT SERIES
16800     170 DO 180 J=1,N
16900         SD(J)=-G(J)+BETA*SD(J)
17000         SD(J+N)=X(J)
17100         DY=DY+SD(J)*G(J)
17200     180 CONTINUE
17300   C  IS SEARCH DIRECTION SUFFICIENTLY DOWNHILL?
17400   C  STEEPEST DESCENT STEP IF NOT
17500         IF(-1.2*GNRM.GT.DY)GO TO 30
17600         IF(DY.GT.-0.8*GNRM)GO TO 30
17700   C
17800   C****************** LINEAR SEARCH ROUTINE **********************
17900   C
18000     200 KSZ=0
18100         ICG(KOUNT,2)=II
18200         ICG(KOUNT,3)=RS
18300         SZ=0.
18400         FI=F
18500         LSZ=0.
18600         DYI=XLS*ABS(DY)
18700         DO 210 J=1,N
18800         SZ=SZ+ABS(SD(J))
18900         GL(J)=G(J)
19000     210 CONTINUE
19100         SZ=(1./SZ)*SFACT
19200         SZI=SZ
19300   C
19400   C  TAKE A STEP ALONG SEARCH DIRECTION
19500     215 DO 220 J=1,N
19600         X(J)=SD(J+N)+SZ*SD(J)
19700     220 CONTINUE
19800   C  COMPUTE FUNCTION VALUE & GRADIENT FOR NEW ARGUMENT
19900         CALL FUNCT(X,F,G,NFE)
20000         IF(SZ.NE.SZI)GO TO 225
20100     225 DY=0.
20200         DO 230 J=1,N
20300         DY=DY+G(J)*SD(J)
20400     230 CONTINUE
20500         IF(DY.LT.0.)GO TO 240
20600   C  MINIMUM PASSED. DY SMALL ENOUGH?
20700         IF(ABS(DY).GT.DYI)GO TO 250
20800         IF(F.GE.FI)GO TO 250
20900         IF(F.GE.FOLD)GO TO 250
21000         ICG(KOUNT,4)=1
21100         FCG(1,KOUNT+1)=F
21200         FCG(2,KOUNT+1)=SZ
21300         RS=0
21400         GO TO 440
21500   C  MINIMUM NOT YET PASSED
21600     240 LSZ=SZ
21700         SZ=2.*SZ
21800         GO TO 215
21900   C
22000   C  MINIMUM HAS BEEN BRACKETED.  SEARCH THIS REGION MORE CLOSELY
22100     250 HSZ=SZ
22200     260 A=(HSZ-LSZ)/2.
22300         IF(ABS(A).GT.ABS(SZR*SZ))GO TO 300
22400   C  STEP SIZE INCREMENT TOO SMALL.
```

```
22500          ICG(KOUNT,4)=2
22600          FCG(1,KOUNT+1)=F
22700          FCG(2,KOUNT+1)=SZ
22800          RS=0
22900          GO TO 340
23000    C
23100    C  EXAMINE ANOTHER POINT IN REGION
23200      300 SZ=LSZ+A
23300          DO 310 J=1,N
23400          X(J)=SD(J+N)+SZ*SD(J)
23500      310 CONTINUE
23600          CALL FUNCT(X,F,G,NFE)
23700          DY=0.
23800          DO 320 J=1,N
23900          DY=DY+G(J)*SD(J)
24000      320 CONTINUE
24100    C  HAS FUNCTION VALUE DECREASED BELOW PREVIOUS MINIMUM VALUE?
24200          IF(ABS(DY).GT.DYI)GO TO 330
24300          IF(F.GE.FI)GO TO 330
24400          IF(F.GE.FOLD)GO TO 330
24500          ICG(KOUNT,4)=3
24600          FCG(1,KOUNT+1)=F
24700          FCG(2,KOUNT+1)=SZ
24800          RS=0
24900          GO TO 440
25000      330 IF(DY.GT.0.)GO TO 250
25100    C  INCREASE STEP SIZE. DY -VE
25200          LSZ=SZ
25300          GO TO 260
25400    C
25500    C******************** CHECK FOR CONVERGENCE ********************
25600    C  COME HERE IF STEPSIZE INCREMENT BECAME TOO SMALL BEFORE F<FI AND
25700    C  F<FOLD WITH ABS(DY)<DYI
25800    C  IF F HAS NOT DECREASED OVER THIS ITERATION, CHECK WITH A SMALLER
25900       STEP SIZE
26000      340 IF(F.GT.FI)GO TO 350
26100    C  THIS ITERATION WAS O.K.
26200          IF(F.LT.FOLD)GO TO 440
26300          GO TO 420
26400    C  THIS ITERATION FAILED.  TRY WITH A SMALLER STEPSIZE.
26500      350 IF(KSZ.EQ.1)GO TO 400
26600          KSZ=1
26700          SZBEST=SZ
26800          SZ=SZI/100.
26900      360 IF(SZ.GT.SZR*SZI)GO TO 380
27000    C  STEPSIZE IS TOO SMALL
27100          DO 370 J=1,N
27200          X(J)=SD(J+N)+SZBEST*SD(J)
27300      370 CONTINUE
27400          CALL FUNCT(X,F,G,NFE)
27500          ICG(KOUNT,1)=5
27600          FCG(1,KOUNT+1)=F
27700          FCG(2,KOUNT+1)=SZBEST
27800          GO TO 400
27900      380 DO 390 J=1,N
28000          X(J)=SD(J+N)+SZ*SD(J)
28100      390 CONTINUE
28200          CALL FUNCT(X,F,G,NFE)
```

```
28300          LSZ=0.
28400          IF(F.LT.FI)GO TO 225
28500          SZ=SZ/10.
28600          GO TO 360
28700    C  IS FI>F ? I.E. DID F DECREASE LAST ITERATION - FAILED THIS ITERATION
28800      400 RS=1
28900          IF(FI.GT.FOLD)GO TO 420
29000    C  LAST ITERATION O.K., THIS ONE FAILED. UPDATE FOLD
29100          FOLD=FI
29200          GO TO 520
29300    C  RESTORE FOLD AND STOP AS EITHER F DID NOT DECREASE THIS ITERATION AND
29400    C  ON THE LAST OR F>FOLD.
29500      420 DO 430 I=1,N
29600          X(I)=XOLD(I)
29700      430 CONTINUE
29800          CALL FUNCT(X,F,G,NFE)
29900          ICG(KOUNT,4)=4
30000          ICG(KOUNT,1)=NFE
30100          FCG(1,KOUNT+1)=F
30200          FCG(2,KOUNT+1)=0.
30300          RETURN
30400    C
30500    C*************************************************************************
30600    C  F HAS DECREASED BELOW FI AND FOLD WITH ABS(DY)<DYI I.E. THINGS ARE
30700    C  GOING WELL, BUT CAN WE STOP YET, PLEASE?
30800    C
30900      440 FOLD=FI
31000          DO 450 J=1,N
31100          XOLD(J)=X(J)
31200      450 CONTINUE
31300    C  COMPUTE CHANGE IN ARGUMENT
31400          DO 470 J=1,N
31500          K=J+N
31600          SD(K)=X(J)-SD(K)
31700          IF(ABS(SD(K))-EPSX)470,470,520
31800      470 CONTINUE
31900    C  TEST GRADIENTS FOR CONVERGENCE
32000          DO 480 J=1,N
32100          IF(ABS(G(J))-EPSG)480,480,520
32200      480 CONTINUE
32300    C
32400    C  CONVERGED.  TERMINATE ONLY IF THIS IS A STEEPEST DESCENT STEP
32500          IF(ICG(KOUNT,2)*ICG(KOUNT,3).NE.1)GO TO 500
32600          IER=0
32700          ICG(KOUNT,1)=NFE
32800          RETURN
32900    C
33000    C  PERFORM A STEEPEST DESCENT STEP TO ENSURE CONVERGENCE, UNLESS LIMIT
33100    C  ITERATIONS HAVE BEEN REACHED.
33200      500 ICG(KOUNT,1)=NFE
33300          IF(KOUNT.LT.LIMIT)GO TO 510
33400          IER=1
33500          RETURN
33600      510 II=1
33700          RS=1
33800          GO TO 10
33900    C
34000    C  NOT YET CONVERGED.  PERFORM ANOTHER ITERATION IF LIMIT ITERATIONS
34100    C  NOT REACHED.
34200      520 ICG(KOUNT,1)=NFE
```

```
34300          IF(KOUNT.LT.LIMIT)GO TO 10
34400          IER=1
34500          RETURN
34600          END
```

APPENDIX II

FORMULATION OF L.Q.G. PROBLEM EQUATIONS

State Equation: $$x_{i+1} = x_i + I_i - u_i \tag{II.1}$$

Cost Function:

$$J = E\left\{\frac{1}{2}(x_N-y_N)S_N(x_N-y_N) + \frac{1}{2}\sum_{i=0}^{N-1}\left[(x_i-y_i)\right.\right.$$

$$\left.\left.A_i(x_i-y_i) + (d_i-u_i)B_i(d_i-u_i)\right]\right\} \tag{II.2}$$

where x_i = reservoir storage

I_i = total inflow

u_i = release

d_i = demand

y_i = (Max. Storage-Min Storage)/2

N = final time.

So A_i penalises deviation from y_i

B_i acts as a quadratic fuel cost parameter.

Now split stochastic quantities into separate mean (e.g. $\overset{n}{u}_i$) and variable ($\hat{u}_i$) components. This can be done because the system is linear and Gaussian. Define:

$$\hat{u}_i + \tilde{u}_i = d_i - u_i - \hat{T}_i - \tilde{T}_i$$

where $\hat{T}_i$, $\tilde{T}_i$ are mean and variable components of tributary flows.

$\hat{u}_i$, $\tilde{u}_2$ are now the mean and variable components of thermal generation.

Substituting in eqn (I.1), for u_i and adding $y_i=y_{i+1}$ to each side:

$$\hat{x}_{i+1} + \tilde{x}_{i+1} + y_{i+1} = \hat{x}_i + \tilde{x}_i + y_i + \hat{I}_i + \tilde{I}_i - d_i + \hat{u}_i$$

$$+ \tilde{u}_i + \hat{T}_i + \tilde{T}_i \tag{II.3}$$

$$\Rightarrow \quad \tilde{x}_{i+1} = \tilde{x}_i + \tilde{I}_i + \tilde{u}_i + \tilde{T}_i \tag{II.4}$$

and $$\hat{x}_{i+1} = \hat{x}_i + \hat{I}_i + d_i + y_i - y_{I+1} + \hat{u}_i + \hat{T}_i$$

$$= \hat{x}_i + \hat{u}_i + Z_i \tag{II.5}$$

where $Z_i = y_i - y_{i+1} - d_i + \hat{I}_i + \hat{T}_i$

So Z_i is simply the sum of the deterministic quantities, for convenience. Note that it has been assumed that tributary and controllable flows are perfectly correlated so $\tilde{T}_i$ will be added to $\tilde{I}_i$ from now on.

If $\alpha_i = \rho_i \, \sigma_i / \sigma_{i-1}$

where ρ_i = serial correlation coefficient (lag one)

Then
$$\tilde{I}_i = \alpha_i \tilde{I}_{i-1} + \beta_i \tag{II.6}$$

(from eqn 7.1).

In matrix form the two state equations are now:

(i) Deterministic

$$\hat{X}_{i+1} = \begin{bmatrix} \hat{x}_{i+1} \\ Z_{i+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & Z_{i+1}/Z_i \end{bmatrix} \begin{bmatrix} \hat{x}_i \\ Z_i \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \hat{u}_i \tag{II.7}$$

$$= \hat{\Phi}_i \qquad = \hat{\Gamma}_i$$

where $\hat{X}_o$ is given

(ii) Stochastic

$$\tilde{X}_{i+1} = \begin{bmatrix} \tilde{x}_{i+1} \\ \tilde{I}_i \end{bmatrix} = \begin{bmatrix} 1 & \alpha_i \\ 0 & \alpha_i \end{bmatrix} \begin{bmatrix} \tilde{x}_i \\ \tilde{I}_{i-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tilde{u}_i + \begin{bmatrix} \beta_i \\ \beta_i \end{bmatrix} \tag{II.8}$$

$$= \tilde{\Phi}_i \qquad = \tilde{\Gamma}_i \qquad = w_i$$

where $\tilde{X}_o = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

The cost function is now:

$$J = \frac{1}{2} E \left\{ X_N^T S_N X_N + \sum_{i=0}^{N-1} \left[X_i^T A X_i + U_i^T B U_i \right] \right\} \qquad \text{(II.9)}$$

where X_N, X_i, U_i are deterministic for calculation of $\overset{n}{J}$, stochastic for $\tilde{J}$

$$J_{Total} = \overset{n}{J} + \tilde{J}$$

$$A = \begin{bmatrix} p & 0 \\ 0 & 0 \end{bmatrix} = \text{penalty on deviations from mean state}$$

$$B = K \quad = \text{thermal fuel cost parameter}$$

$$S_N = \begin{bmatrix} p_N & 0 \\ 0 & 0 \end{bmatrix}$$

These two problems can now be solved in the usual way, see for example Bryson and Ho (1975). The stochastic and deterministic components both have solutions of the same form i.e.

$$u_i = -C_{i,1}\, x_{i,1} - C_{i,2}\, x_{i,2}$$

$$= -C_i\, x_i \qquad \text{(II.10)}$$

where $C_i = (\Gamma^T S_{i+1} \Gamma + B)^{-1} (\Gamma^T S_{i+1} \Phi_i)$ (II.11)

and $S_i = \Phi_i^T S_{i+1} \Phi_i - C_i^T (B + \Gamma^T S_{i+1} \Gamma) C_i + A$ (II.12)

$\tilde{X}_{i+1}$ can be rewritten, eliminating $\tilde{u}_i$:

$$\tilde{X}_{i+1} = (\tilde{\Phi}_i - \tilde{C}_i)\, \tilde{X}_i + w_i \qquad \text{(II.13)}$$

The state covariance matrix, P_{i+1}, can now be found:

$$P_{i+1} = E\left[\tilde{X}_{i+1}, \tilde{X}_{i+1}\right]$$

$$= (\tilde{\Phi}_i - \tilde{C}_i) P_i (\tilde{\Phi}_i - \tilde{C}_i)^T + w_i\, w_i^T \qquad \text{(II.14)}$$

where $P_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ as there is no uncertainty in the initial state.

$P_{i\ 1,1} = \sigma^2_{xi}$, variance of storage

$P_{i\ 2,2} = \sigma^2_{I\ i-1}$ variance of previous inflow.

Results for this model appear in section 7.4

APPENDIX III

INTEGRATION DETAILS FOR GAUSSIAN PLUS IMPULSE MODEL

For calculation of $\mu(t+1)$:

$$\int_{-\infty}^{X_{max}} \frac{Z}{\sigma(t)\sqrt{2\pi}} e^{-(Z-\mu(t))^2/2\sigma(t)^2} \, dZ$$

$$= \int_{-\infty}^{X_{max}} \left(\frac{(Z-\mu(t))\sigma(t)}{\sigma^2(t)\sqrt{2\pi}} + \frac{\mu(t)}{\sigma(t)\sqrt{2\pi}} \right) e^{-(Z-\mu(t))^2/2\sigma(t)^2} \, dZ$$

$$= \frac{\sigma(t)}{\sqrt{2\pi}} \left[-e^{-(Z-\mu(t))^2/2\sigma(t)^2} \right]_{-\infty}^{X_{max}} + \frac{\mu(t)}{2} \left(1 + \text{erf} \left(\frac{X_{max}-\mu(t)}{\sigma(t)\sqrt{2}} \right) \right)$$

$$= \frac{-\sigma(t)}{\sqrt{2\pi}} e^{-(X_{max}-\mu(t))^2/2\sigma(t)^2} + \frac{\mu(t)}{2} \left(1 + \text{erf} \left(\frac{X_{max}-\mu(t)}{\sigma(t)\sqrt{2}} \right) \right)$$

For calculation of $\sigma^2(t+1)$:

$$\int_{-\infty}^{X_{max}} \frac{Z^2}{\sigma(t)\sqrt{2\pi}} e^{-(Z-\mu(t))^2/2\sigma(t)^2} \, dZ$$

$$= \frac{2\sigma(t)}{\sqrt{2\pi}} \int_{-\infty}^{X_{max}} \frac{Z^2}{2\sigma(t)^2} e^{-(Z-\mu(t))^2/2\sigma(t)^2} \, dZ$$

$$= \frac{2\sigma(t)}{\sqrt{2\pi}} \int_{-\infty}^{X_{max}} \left(\frac{(Z-\mu(t))^2}{2\sigma(t)^2} + \frac{2Z\mu(t)-\mu(t)^2}{2\sigma(t)^2} \right) e^{-\frac{(Z-\mu(t))^2}{2\sigma(t)^2}} \, dZ$$

This is integrated as three components:

(i) $$I_1 = \int_{-\infty}^{\mu(t)} \frac{(Z-\mu(t))^2}{2\sigma(t)^2} e^{-(Z-\mu(t))^2/2\sigma(t)^2} \, dZ$$

$$= \int_{h=-\infty}^{h=0} h^2 e^{-h^2} \sqrt{2}\sigma \, dh$$

by making the substitution $h = \frac{Z-\mu(t)}{\sqrt{2}\,\sigma(t)}$

$$\Rightarrow dZ = \sqrt{2}\,\sigma(t)\,dh$$

$$\therefore \quad I_1 = \sqrt{2}\,\sigma \cdot \frac{\Gamma\left[3/2\right]}{2}$$

using an integral from Spiegel (1968), where Γ = the Gamma function

$$I_1 = \frac{\sigma(t)\sqrt{2\pi}}{4}$$

(ii) $$I_2 = \int_{\mu(t)}^{X_{max}} \frac{(Z-\mu(t))^2}{2\,\sigma(t)^2}\, e^{-(Z-\mu(t))^2/2\sigma(t)^2}\, dZ$$

$$= -\left[\frac{(Z-\mu(t))}{2}\, e^{-\frac{(Z-\mu(t))^2}{2\,\sigma(t)^2}}\right]_{\mu(t)}^{X_{max}} + \int_{\mu(t)}^{X_{max}} \frac{1}{2}\, e^{-(Z-\mu(t))^2/2\sigma(t)^2}\, dZ$$

$$= -\frac{X_{max}-\mu(t)}{2}\, e^{-\frac{(X_{max}-\mu(t))^2}{2\sigma(t)^2}} + \frac{\sigma(t)\sqrt{2\pi}}{4}\, \mathrm{erf}\left[\frac{X_{max}-\mu(t)}{\sigma(t)\sqrt{2}}\right]$$

using integration by parts:

$$\int fg' = fg - \int gf'$$

where $f = \frac{Z-\mu(t)}{2}$ $\quad f' = \frac{1}{2}$

$$g = -e^{-(Z-\mu(t))^2/2\sigma(t)^2} \qquad g' = \frac{Z-\mu(t)}{\sigma(t)^2}\, e^{-\frac{(Z-\mu(t))^2}{2\sigma(t)^2}}$$

(iii) $$I_3 = \int_{-\infty}^{X_{max}} \frac{2Z\mu(t)-\mu(t)^2}{2\sigma(t)^2}\, e^{-(Z-\mu(t))^2/2\sigma(t)^2}\, dZ$$

$$= \int_{-\infty}^{X_{max}} \mu(t) \left[\frac{Z-\mu(t)}{\sigma(t)^2}\right] e^{-\frac{(Z-\mu(t))^2}{2\sigma(t)^2}}\, dZ$$

$$+ \int_{-\infty}^{X_{max}} \frac{\mu(t)^2}{2\sigma(t)^2}\, e^{-(Z-\mu(t))^2/2\sigma(t)^2}\, dZ$$

$$= \left[-\mu\, e^{\frac{-(Z-\mu(t))^2}{2\sigma(t)^2}} \right]_{-\infty}^{X_{max}} + \frac{\mu(t)^2}{2\sigma(t)} \cdot \sqrt{2\pi} \cdot \frac{1}{2} \left(1+\mathrm{erf} \left(\frac{X_{max}-\mu(t)}{\sigma(t)\ \sqrt{2}} \right) \right)$$

$$= -\mu\, e^{-(X_{max}-\mu(t))^2/2\sigma(t)^2} + \frac{\mu^2(t)\sqrt{2\pi}}{4\sigma(t)} \left(1+\mathrm{erf} \left(\frac{X_{max}-\mu(t)}{\sigma(t)\ \sqrt{2}} \right) \right)$$

REFERENCES

Abramowitz, M., and Stegun, I. A., (eds) "Handbook of Mathematical Functions", Dover.

Agarwal, S. K., and Nagrath, I. J. (1972). "Optimal scheduling of hydro-thermal systems", Proc. IEE, 119, pp 169-172.

Alspach, D. L., and Sorenson, H. W. (1972). "Non-linear Bayesian estimation using Gaussian sum approximations", IEEE Trans. Auto. Control AC-17, pp 439-448.

Anderson, B. D. O., and Moore, J. B. (1979). "Optimal Filtering", Prentice-Hall.

Anonymous, (1972). "A New Algorithm for Optimization", Math. Prog., 3, 124-128.

Arvanitis, N. V., and Rosing, J. (1970). "Composite representation of a multi-reservoir hydro-electric power system", IEEE Trans. Power Appar. Syst. PAS-89, pp 319-335.

Beale, E. M. L. (1972). "A derivation of conjugate gradients" in Numerical Methods for Non-Linear Optimisation, F. A. Lootsma ed., pp 39-44, Academic Press.

Beckman, F. S. (1962). "The solution of linear equations by the conjugate gradient method" in Mathematical Methods for Digital Computers, Vol. 1, A. Ralston and H. S. Wilf, eds. pp 62-72.

Bellman, R. (1961). "Adaptive Control Processes - a Guided Tour", Princeton University Press.

Boshier, J. F., and Lermit, R. J. (1977). "A network flow formulation for optimum reservoir management of the New Zealand power generating system", N.Z.O.R., 5, pp 85-100.

Bryson, A. E., and Ho, Y. C. (1975). "Applied Optimal Control", Halsted.

Chikhani, A. Y., Quintana, V. H., and Chan, P. T. L. (1979). "A Stochastic approach for scheduling of hydro units in a hydro-thermal system", 1979 Control of Power Systems Conf. and Exposition, Conf. Record, pp 61-67.

Daellenbach, H. G. (1979). "Long term scheduling of hydro-thermal power systems: D. P. formulation with decomposition by prices", Proc. ORSNZ Conf., pp 33-39.

Daellenbach, H. G., and Read, E. G. (1976). "Survey on optimisation of the long-term scheduling of hydro-thermal power systems", presented to Annual Conf. ORSNZ, Auckland.

Dillon, T. S. (1974). "Problems of optimal economic operation and control of integrated and thermal power systems", Ph.D. Thesis, Monash University, Australia.

Dillon, T. S., Martin, R. W., and Sjelvgen, D. (1980). "Stochastic optimisation and modelling of large hydro-thermal systems for long-term regulation", Electrical Power & Energy Systems, 2, pp 2-20.

Dillon, T. S., and Morsztyn, K. (1972). "A New theoretical and computational approach to the exact solution of the problem of the optimal control of integrated (hydro-thermal) power systems", 5th I.F.A.C. World Congress.

El-Hawary, M. E., and Christensen, G. S. (1979). "Optimal economic operation of power systems", Academic Press.

Fletcher, R., and Reeves, C. M. (1964). "Function minimisation by conjugate gradients", Computer Journal, 7, pp 149-154.

Fults, D. M., Hancock, L. F., and Logan, G. R. (1976). "A Practical monthly optimum operations model", J. Water Resources Planning and Management Division, Proc. ASCE, 102, pp 63-76.

Gagnon, C. R., Hicks, R. H., Jacoby, S. L. S. and Kowalik, J. S. (1974). "A Non-linear programming approach to a very large hydro-electric system optimisation", Math. Prog., 6, pp 28-41.

Green, D. J. (1971a). "Optimal control of a cascade of hydro-electric stations", N.Z.O.R., 1, pp 21-32.

Green, D. J. (1971b). "The Optimal operation of hydro-electric stations on a single river", M.E. Thesis, University of Canterbury, New Zealand.

Hanscom, M. A., Lafond, L., Lasdon, L., and Pronovost, G. (1980). "Modelling and resolution of the medium term energy generation planning problem for a large hydro-electric system", Management Science, 26, pp 659-668.

Hicks, R. H., Gagnon, C. R., Jacoby, S. L. S., and Kowalik, J. S. (1974). "Large scale non-linear optimisation of energy capability for the Pacific Northwest hydro-electric system", IEEE Trans. Power Appar. Syst. PAS-93, pp 1604-1612.

Kirchmayer, L. K. (1958). "Economic Operation of Power Systems", Wiley.

Kuester, J. L. and Mize, J. H. (1973). "Optimisation Techniques with Fortran", McGraw-Hill.

Kumar, S., Sharma, J., and Ray, L. M. (1979). "A Non-linear programming algorithm for hydro-thermal generation scheduling", Computers & Elect. Eng., 6, pp 221-229.

Lasdon, L. S. (1970). "Optimisation theory for large systems", MacMillan.

Linsley, R. K., Kohler, M. A. and Paulhus, J. L. H. (1958). "Hydrology", for Engineers", McGraw-Hill.

Lusk, T. A. (1972). Management of hydro reservoirs in the New Zealand power system", Notes used for NZIE/ORSOC meeting, Christchurch.

Maidment, D. R. (1976). "Stochastic state space approach to reservoir control", Presented at IIASA/WMO workshop on the recent developments in real time forecasting/control of water resource systems, Laxenburg, Austria.

Matalas, N. C. (1967). "Time Series Analysis", J. Water Resources Research, 3, pp 817-829.

McCool, K. D., Howard, R. K., and Couch, R. L. (1966). "A computer programme for simulated operation of the New Zealand power system", unpublished New Zealand Electricity report.

McKerchar, A. I. (1971). "Optimal operation of water resource systems", Ph.D. Thesis, University of Canterbury, New Zealand.

McKerchar, A. I. (1975). "Optimal monthly operation of inter-connected hydro-electric power storages", Journal of Hydrology, 25, pp 137-158.

McPike, A. W. (1981). "Lake Waikaremoana: Methods of investigating leakage through the natural dam", NZIE Conf., Auckland.

Patchell, J. W. and Jacobs, O. L. R. (1971). "Separability, neutrality and certainty equivalence", Int. J. Control, 13, pp 337-342.

Peters, R. J., Chu, K-C., and Jamshidi, M. (1978). Optimal operation of a water resources system by stochastic programming", Mathematical Programming Study 9, North Holland.

Podmore, R. (1972). "Digital Computer Analysis of Power System Networks", Ph.D. Thesis, University of Canterbury, New Zealand.

Podmore, R. (1973). "A Simplified and improved method for calculating transmission loss formulas", Proc. of 8th P.I.C.A. Conf., pp 433-440.

Podmore, R. (1974). "Economic power dispatch with line security limits", IEEE Trans. Power Appar. Syst. PAS 93 pp 289-295.

Polak, E., and Ribiere, G. (1969). "Note sur la convergence de methods de directions conjugres", Revue Francaise Informat. Recherce Operationnelle, 16, pp 35-43.

Powell, M. J. D. (1977). "Restart Procedures for the Conjugate gradient method", Math. Prog., 12, pp 241-254.

Quintana, V. H., and Chikhani, A. Y. (1981). "A Stochastic model for mid-term operation planning of hydro-thermal systems with random reservoir inflows", IEEE Trans. Power Appor. Syst. PAS 100, pp 1119-1127.

Quintana, V. H., Chikhani, A. Y., and Chan, P. T. L. (1979). "Stochastic simulation of the long-term operation of hydro-thermal systems with random reservoir inflows", Proc. I.F.A.C. Conf., Calcutta, Vol. III, pp 40-48.

Read, E. G. (1979). "Optimal Operation of power systems", Ph.D. Thesis, University of Canterbury, New Zealand.

Revelle, C., and Gundelach, J. (1975). "The Linear decision rule in reservoir management and design 4. A rule that minimises output variance", Water Resources Research, 11, pp 197-203.

Revelle, C., Joeres, E., and Kirby, W. (1969). "The Linear decision rule in reservoir management and design 1. Development of the model", Water Resources Research, 5, pp 767-777.

Rockafellar, R. T., and Wets, R. J-B. (1976). "Nonanticipativity and L' martingales in stochastic optimisation problems", in R. J-B. Wets ed, Mathematical Programming Study 6, Vol. II, North Holland, pp 170-187.

Rosenthal, R. E. (1980). "The Status of optimisation models for the operation of multi-reservoir systems with stochastic inflows and non-separable benefits", Research Report No. 75, Water Resources Research Centre, The University of Tennessee.

Rosenthal, R. E. (1981). "A Non-linear network flow algorithm for maximisation of benefits in a hydro-electric power system", Operations Research, 29, 763-787.

Sachdeva, S. S. (1982). "Bibliography on optimal reservoir drawdown for the hydro-electric power system", IEEE Trans. Power Appor. Syst. PAS 101, pp 1482-1496.

Saha, T. N., and Khaparde, S. A. (1978). "An Application of a direct method to the optimal scheduling of a hydro-thermal system", IEEE Trans. Power Appar. Syst. PAS 97, pp 977-983.

Shanno, D. F. (1978). "Conjugate gradient methods with inexact searches", Mathematics of Operations Research, 3 pp 244-256.

Soares, S., Lyra, C., and Tavores, H. (1980). "Optimal Generation scheduling of hydro-thermal power system", IEEE Trans Power Appar. Syst. PAS 99, pp 1107-1118.

Sorenson, H. W. (1976). "An Introduction to Non-linear programming -" Computers and Electrical Eng., Vol. 3:
"I Necessary and Sufficient Conditions" pp 3-34.
"II The Linear Programming Problem", pp 127-157.
"III Search Procedures for Unconstrained Minimisation" pp 239-269.
"IV Numerical Methods for Constrained Minimisation" pp 347-386.

Sorenson, H. W., and Alspach, D. L. (1971). "Recursive Bayesian estimation using Gaussian sums", Automatica, 7, pp 465-479.

Spiegel, M. R. (1968). "Mathematical Handbook of Formulas and Tables", McGraw-Hill.

Stage, S., and Larson, Y. (1961). "Incremental cost of water power", A.I.E.E. Trans. Pt III, 80, pp 361-365.

Stroud, A. H., and Secrest, D. (1966). "Gaussian Quadrature Formulas", Prentice-Hall.

Vajda, S. (1972). "Probabilistic Programming", Academic Press.

Viramontes, F. A., and Hamilton, H. B. (1978). "Optimal long range hydro scheduling in the integrated power system", IEEE Trans. Power Appar. Syst. PAS 97, pp 292-299.

Yevjevich, V. (1972). "Stochastic Processes in Hydrology", Water Resources Publications, Fort Collins, Colorado.